Molecular Mechanisms of Plasmodium Red Blood Cell Invasion

Brian Michael Malpede
Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds

Part of the Biology Commons

Recommended Citation
https://openscholarship.wustl.edu/art_sci_etds/456

This Dissertation is brought to you for free and open access by the Arts & Sciences at Washington University Open Scholarship. It has been accepted for inclusion in Arts & Sciences Electronic Theses and Dissertations by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.
Molecular Mechanisms of *Plasmodium* Red Blood Cell Invasion

by

Brian Michael Malpede

A dissertation presented to the Graduate School of Arts & Sciences of Washington University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

May 2015

St. Louis, Missouri
## Table of Contents

List of Figures ......................................................................................................................... iv
List of Tables ............................................................................................................................... vi
List of Abbreviations ..................................................................................................................... vii
Acknowledgements ....................................................................................................................... viii
Abstract of the Dissertation ........................................................................................................ ix

### Chapter 1: Introduction ........................................................................................................... 1

1.1 Malaria disease and *Plasmodium* species ............................................................................ 3

1.2 The structure and function of *Plasmodium* adhesins .......................................................... 5

1.2.1 Proteins functioning as a parasite surface coat ................................................................... 5

1.2.2 The CSP surface coat on sporozoites ................................................................................ 8

1.2.3 Surface proteins that adhere to cellular receptors during invasion ..................................... 9

1.2.4 Adhesion during membrane invagination to form the parasitophorous vacuole ............... 12

1.2.5 The link between host cell receptors and the parasite's actin motor ............................... 14

1.2.6 Adhesive proteins exported to the infected RBC surface ............................................... 15

1.3 Targeting adhesion proteins in vaccines ............................................................................. 17

1.3.1 RBC invasion ligands as targets for vaccination ............................................................... 17

1.4 Figures ................................................................................................................................. 19

1.5 References ............................................................................................................................ 22

### Chapter 2: Identification of a receptor binding interface and determinants of receptor

specifcity for PfEBA-140 RBC invasion .................................................................................... 31

2.1 Abstract .................................................................................................................................. 34

2.2 Introduction ........................................................................................................................... 35

2.3 Experimental Procedures ..................................................................................................... 37

2.4 Results .................................................................................................................................. 40

2.5 Discussion .............................................................................................................................. 44

2.6 Figures and Tables ................................................................................................................ 47

2.7 References ............................................................................................................................ 52
Chapter 3: The molecular basis for sialic acid dependent receptor recognition by PfEBA-140

3.1 Abstract ........................................................................................................................................57
3.2 Introduction ....................................................................................................................................58
3.3 Experimental Procedures ..............................................................................................................62
3.4 Results ...........................................................................................................................................64
3.5 Discussion .......................................................................................................................................71
3.6 Figures and Tables ..........................................................................................................................75
3.7 References ......................................................................................................................................82

Chapter 4: PvDBP utilizes unique receptor binding interfaces and a novel two-step binding process to engage its RBC receptor .................................................................................................85

4.1 Abstract .........................................................................................................................................88
4.2 Author Summary ..............................................................................................................................89
4.3 Introduction .....................................................................................................................................90
4.4 Results ............................................................................................................................................94
4.5 Discussion ......................................................................................................................................101
4.6 Experimental Procedures .............................................................................................................107
4.7 Figures and Tables ..........................................................................................................................111
4.8 References ....................................................................................................................................126

Chapter 5: Conclusions and Future Directions ..................................................................................130

5.1 Introduction ...................................................................................................................................131
5.2 Receptor engagement mechanisms of the EBL family .................................................................133
5.3 Targeting functional elements of DBL domains for vaccination ....................................................138
5.4 EBL ligands in a multi-component vaccine ......................................................................................141
5.5 Figures ..........................................................................................................................................144
5.6 References .....................................................................................................................................145
List of Figures

Chapter 1: Introduction .......................................................................................................................1
  Figure 1. Domain architectures of adhesive proteins functioning at different parasite life stages ...19
  Figure 2. Crystal structures define adhesive folds used by the malaria parasite ...............20

Chapter 2: Identification of a receptor binding interface and determinants of receptor
specificity for PfEBA-140 RBC invasion ............................................................................................31
  Figure 1. Structure of RII PfEBA-140 ............................................................................................47
  Figure 2. PfEBA-140 erythrocyte binding requires both DBL domains .................................48
  Figure 3. Structural determinants of receptor specificity within the EBL family .................49
  Figure 4. Oligomeric state of RII PfEBA-140 ..............................................................................50

Chapter 3: The molecular basis for sialic acid dependent receptor recognition by PfEBA-
140......................................................................................................................................................54
  Figure 1. The crystal structure of RII PfEBA-140 bound to sialyllactose reveals two
  receptor glycan binding sites ........................................................................................................75
  Figure 2. Sialic acid modeling and erythrocyte binding studies provide a detailed
  molecular description of the RII PfEBA-140 erythrocyte invasion interface ....................76
  Figure 3. RII PfEBA-140 glycan binding sites are distinct from other classes of sialic acid
  binding proteins .............................................................................................................................77
  Figure 4. The F1 binding pocket surface is altered when Ile185 is mutated to valine in
  silico ................................................................................................................................................78
  Figure 5. The sialyllactose bound structure illuminates a glycan induced structural change
  and allows for localization of putative sulfate-binding sites .................................................79
  Figure 6. PfEBA-140 sialic acid recognition is distinct from that of PfEBA-175 .............80

Chapter 4: PvDBP utilizes unique receptor binding interfaces and a novel two-step binding
process to engage its RBC receptor ..................................................................................................85
  Figure 1. Residues 14-43 of DARC contain the minimal binding region ...............................111
  Figure 2. Crystal Structure of the DBP-RII:DARC heterotrimer and heterotetramer .......112
Figure 3. Isothermal titration calorimetry reveals step-wise binding of DARC to DBP-RII in solution.................................................................113
Figure 4. Binding interfaces of the DBP-RII:DARC heterotrimer .................................114
Figure 5. Binding interfaces of the DBP-RII:DARC heterotetramer ...............................115
Figure 6. The structural studies define red blood cell binding ......................................116
Figure 7. Mapping polymorphic residues and inhibitory epitopes reveals targets of selective pressure ........................................................................................................117
Figure 8. A model for attachment during invasion .........................................................118
Figure S1. DARC residues 19-30 are contacted by DBP-RII...........................................119
Figure S2. Upon receptor binding, new regions of DBP-RII become structured, while preexisting structural regions undergo no major conformational changes ..................120
Figure S3. The Sulfotyrosine Binding Site ....................................................................121
Figure S4. The DARC binding pockets are distinct from residues previously suggested to bind DARC from mutagenesis studies..........................................................122

Chapter 5: Conclusions and Future Directions................................................................130

Figure 1. Summary of proposed invasion mechanisms for Plasmodium EBL family members .................................................................................................................144
List of Tables

Chapter 2: Identification of a receptor binding interface and determinants of receptor specificity for PfEBA-140 RBC invasion ..................................................................................31
  Table 1. Data collection and refinement statistics for RII PfEBA-140.........................51

Chapter 3: The molecular basis for sialic acid dependent receptor recognition by PfEBA-140................................................................................................................................................54
  Table 1. Data collection and refinement statistics for PfEBA-140 RII sialyllactose complex.................................................................................................................................81

Chapter 4: PvDBP utilizes unique receptor binding interfaces and a novel two-step binding process to engage its RBC receptor ................................................................................85
  Table 1. Data Collection and Refinement Statistics for DBP-RII:DARC heterotrimer and heterotetramer ......................................................................................................................123
  Table 2. Heterotrimer interface residues determined by PDBePISA ..........................124
  Table 3. Heterotetramer interface residues determined by PDBePISA .........................125
List of Abbreviations

AMA-1 – Apical Membrane Antigen-1
CIDR – Cysteine Rich Interdomain Region
DARC – Duffy Antigen and Receptor for Chemokines
DBL – Duffy Binding-like
EBL – Erythrocyte-binding-like
F1/F2 – Description of the individual DBL domains in *P. falciparum* EBL Region II
GAG – Glycosaminoglycan
GPA – Glycophorin A
GPB – Glycophorin B
GPC – Glycophorin C
ITC – Isothermal Titration Calorimetry
MIC – Micronemal Invasion Protein Family of *Toxoplasma gondii*
NMR – Nuclear Magnetic Resonance
PfEBA-140 – *Plasmodium falciparum* Erythrocyte Binding Antigen 140
PfEBA-175 – *Plasmodium falciparum* Erythrocyte Binding Antigen 175
PfEBA-181 – *Plasmodium falciparum* Erythrocyte Binding Antigen 181
PfEBL-1 – *Plasmodium falciparum* Erythrocyte Binding Antigen 1
PfEMP1 – *Plasmodium falciparum* Erythrocyte Membrane Protein 1
PkDBP – *Plasmodium knowlesi* Duffy Binding Protein
PvDBP – *Plasmodium vivax* Duffy Binding Protein
RBC – Red Blood Cell
RII – Region II
SAXS – Small Angle X-ray Scattering
TRAP – Thrombosponding Related Anonymous Protein
TSR – Thrombospondin Type-I Repeat
VWA – Von-Willebrand Factor Type A
Acknowledgments

I am incredibly grateful to have the infinite support and love of my mom, dad, brother, and partner Victoria. They are all incredible examples of drive, perseverance, and dedication. I am also indebted to the many friends I made in the past five years, without whom graduate school just wouldn’t have been the same. Many thanks to my thesis adviser Niraj for providing me with so many opportunities to succeed, learn, and develop as a scientist and for being such a strong supporter and guiding element for the past four years. And finally, thank you to my committee for providing direction to my research and for such wonderful support as I grew as an independent researcher.
ABSTRACT OF THE DISSERTATION

Molecular Mechanisms of *Plasmodium* Red Blood Cell Invasion

by

Brian Michael Malpede

Doctor of Philosophy in Biology and Biomedical Sciences
Molecular Microbiology and Microbial Pathogenesis
Washington University in St. Louis, 2015
Niraj H. Tolia, Chair

The malaria parasite *Plasmodium* utilizes specialized proteins for adherence to cellular receptors in its mosquito vector and human host. Adherence is critical for parasite development, host cell traversal and invasion, and protection from vector and host immune mechanisms. These vital roles have identified several adhesins as vaccine candidates. A deficiency in current adhesin-based vaccines is induction of antibodies targeting non-conserved, non-functional, and decoy epitopes due to the use of full length proteins or binding domains. To alleviate the elicitation of non-inhibitory antibodies, conserved functional regions of proteins must be identified and exploited. Structural biology provides the tools necessary to achieve this goal, and has succeeded in defining biologically functional receptor binding and oligomerization interfaces for a number of promising malaria vaccine candidates.

A critical step in the *Plasmodium* life cycle is adhesion to and invasion of erythrocytes by the merozoite during blood stage growth. The invasion process requires the formation of a tight junction between parasite adhesive proteins and red cell receptors. Formation of this junction allows the merozoite to engage and enter the red cell to form the parasitophorous vacuole. The
Erythrocyte-binding-like (EBL) family of cellular adhesion proteins is implicated in the formation of this essential tight junction.

Receptor binding by the EBL family of invasion ligands has been localized to a specific portion of each EBL protein designated Region II (RII). This region is composed of the Duffy-binding-like (DBL) domain, unique to *Plasmodium* species. The two *Plasmodium* species responsible for the majority of malaria disease in humans are *P. falciparum* and *P. vivax*. There are four identified EBL members in *P. falciparum* that contain two tandem DBL domains in RII. This organization contrasts the single DBL domain in RII of the only characterized EBL member of *P. vivax*, Duffy-binding Protein (PvDBP). Using a combination of structural, biophysical, and cellular interaction methods we examined the binding mode for these two Plasmodium EBL family members to define the molecular and mechanistic basis of unique red cell invasion routes.

Specifically, we sought to enhance our understanding of the molecular and mechanistic details of invasion by *Plasmodium* EBL ligands with the ultimate goal of identifying novel targets for rational vaccination efforts. We examined receptor interactions of a *P. falciparum* EBL member, Erythrocyte-binding-antigen 140 (PfEBA-140) and *P. vivax* DBP to provide insight into invasion by both species of parasite. Our work identified unique interfaces on the DBL domain that provide specific recognition of each ligand’s unique red blood cell receptor. Additionally, studies examining oligomeric state identified distinct mechanistic parameters for the two EBL ligands, and contrasted with the mechanism observed for another *P. falciparum* member, PfEBA-175.
Chapter 1

Introduction
Preface

The thesis introduction was written and figures were prepared by me with comments from Niraj H. Tolia. A portion of this chapter was published as a review: [Malpede, B.M. and Tolia, N.H. (2014) Malaria Adhesins: Structure and Function. Cellular Microbiology. DOI: 10.1111/cmi.12276].
1.1 Malaria disease and *Plasmodium* species

Malaria eradication efforts have reduced the burden of this devastating parasitic disease, however, 3.2 billion people are still at risk for infection, and approximately 198 million cases of malaria occur annually and result in an estimated 584,000 deaths (WHO World Malaria Report 2014). The disease disproportionately affects the sub-Saharan region of Africa, and children under the age of 5 years account for 78% of all mortality (WHO World Malaria Report 2014).

The disease is the result of infection with one of five species of *Plasmodium*, a genus of protozoan parasites. These parasites engage in a complex life cycle involving growth in a vector of specific species of anopheles mosquito as well as the human host. Human infection is initiated during the blood meal of the female anopheles mosquito, after which the parasite traverses through the human dermis, making its way to the liver to replicate until merozoites are released into the blood. Merozoites travelling the bloodstream then invade and develop within red blood cells (RBCs). This expands the merozoite population and also allows for the development and release of gametes for vector uptake and initiation of a new transmission cycle. The clinical symptoms and death associated with malaria are the result of intricate host-pathogen interactions during blood stage growth and are dependent on the parasite attaching to and invading RBCs (Cowman et al., 2006, Cowman et al., 2012, Miller et al., 2013). Cellular adhesion is essential during growth in both the mosquito vector and human host, and is a multifaceted physical ability that affords the parasite access to unique cellular niches, aids in parasite reproduction, and provides protection from host immunity.

Specifically, adhesion of *Plasmodium* malaria parasites to host cells is critical in mediating traversal through cellular barriers, cellular invasion, and protection from host clearance. To traverse host cells, the parasite disrupts the host membrane, glides through the cytosol, and exits
the cell (Mota et al., 2001). This movement contrasts cellular invasion, during which the parasite engages the host cell and invaginates the membrane to form the parasitophorous vacuole, where the parasite resides as it develops internally (Baum et al., 2008, Cowman et al., 2012). Adhesins also mediate rosetting and cytoadherence of infected erythrocytes within the human host, effectively providing protection from immune clearance by the spleen (Rowe et al., 2009).
1.2 The structure and function of *Plasmodium* adhesins

Proteins that provide adhesive and invasion functions contain defined domains and are organized into families based on domain similarity. Atomic resolution crystal structures of adhesins, alone and in complex with host receptors, have delineated protein folds and structurally conserved segments. More importantly, structures have identified receptor binding interfaces, multimeric contacts, and mechanisms of receptor binding. Evaluation of protein ultrastructure and oligomeric state complemented crystallographic studies by illuminating an overall view of protein shape, flexibility, and multimeric assembly in solution.

Described below is our current knowledge of how protein structures of malaria adhesins provide wide ranging cell adherence functions with defined roles in the parasite life cycle. Families of proteins that are proposed to coat the parasite at multiple life cycle stages are discussed first, followed by the multiple adhesive roles of the sporozoite surface coat protein. Proteins the parasite utilizes to invade RBCs are then detailed. The role of adhesion throughout parasite invagination of the host membrane during cellular invasion is discussed, along with proteins that provide a link between extracellular receptors and the parasite’s internal actin motor. In addition, I provide a discussion of adhesins exported to the surface of infected RBCs that bind numerous receptors to provide rosetting and cytoadherence capabilities. I link the common structural features that serve similar adhesive purposes (Figure 1), and highlight multimeric assembly necessary for adhesion.

**Proteins functioning as a parasite surface coat**

The mosquito blood meal initiates the growth of the parasite in its vector, as gametes enter the proper environment for fertilization. Surface coat adhesins are immediately required, and gametes utilize members of the conserved Apicomplexan 6-cysteine family for recognition and
attachment during fertilization (van Dijk et al., 2001, van Dijk et al., 2010). 6-cysteine family members are also expressed on the surface of sporozoites and merozoites and are likely redundant adhesins at cellular invasion interfaces (Ishino et al., 2005, Sanders et al., 2005). Members in this family contain a varied number of s48/45 domains, a predominantly beta-sheet fold (Fig. 2A) (Arredondo et al., 2012, Tonkin et al., 2013). The crystal structure of Pf12, a merozoite protein, illuminated only minor contacts between the tandem s48/45 domains, and showed that the domain linker lacks conformationally restrictive residues (Tonkin et al., 2013). This suggests that movement between two tandem s48/45 domains is possible, and thus that the link between the two domains is flexible. Furthermore, Pf12 forms heterodimers with another 6-cysteine member, Pf41, suggesting that 6-cysteine family proteins may function as pairs (Taechalertpaisarn et al., 2012).

The domain fold of 6-cysteine family members contrasts the fold of the surface coat proteins on the fertilized zygote and ookinete within the mosquito midgut. After fertilization, zygotes are capable of adhering to one another through tubule extensions coated in the protein p25, an adhesin that contains four tandem evolutionarily conserved epidermal growth factor-like (EGF) domains (Fig. 2B) (Saxena et al., 2006, Rupp et al., 2011). p25 and its homolog, p28, are subsequently expressed on the surface of the motile ookinete, the parasite form that results from the developing zygote (Saxena et al., 2007). As ookinete surface proteins, p25 and p28 are thought to provide protection from mosquito proteolytic defense mechanisms and to facilitate adhesion to the midgut membrane (Tomas et al., 2001, Saxena et al., 2007). Adherence likely depends on p25/p28 binding to laminin, which constitutes a large portion of the midgut epithelium (Vlachou et al., 2001). Knocking out both p25 and p28 is required to severely limit
midgut crossing, suggesting both proteins are critical for adhesion and are functionally redundant (Tomas et al., 2001).

Crystallization of *P. vivax* p25 identified extensive contacts between monomers in the crystal lattice (Fig. 2B) (Saxena et al., 2006). Contacts between p25 monomers form a “triangular prism” that may link the proteins on the ookinete surface to form a specific coat structure (Tomas et al., 2001, Saxena et al., 2006, Saxena et al., 2007). It is important to note that the way in which a protein packs into a crystal can result in artificial contacts that may not be functional *in vivo*. Crystal packing interfaces and oligomeric states should be assessed in solution to support inferences made from crystal structures. The intermolecular contact residues are conserved among p25 orthologs and monomers can self-interact in solution, supporting a physiological role for multimerization (Siden-Kiamos et al., 2000, Saxena et al., 2006, Saxena et al., 2007). Together, these studies suggest that multimeric complexes of p25 and/or p28 form a protective coat and initiate and maintain interactions with the midgut cell wall (Tomas et al., 2001).

In contrast to the surface coat of the ookinete, the sporozoite does not express any characterized EGF containing proteins, but instead utilizes circumsporozoite protein (CSP), discussed below. However, EGF-like domains are used again on the surface of the merozoite after release from the human liver. The merozoite surface protein family (MSP), of which MSP1 is the most abundant, coats the merozoite and provides adhesive functions. MSP1 is a 185–215 kDa protein that undergoes extensive proteolytic processing resulting in multiple subunits non-covalently linked to the C-terminal portion designated MSP1-19 (Fig. 2C). At the point of RBC invasion, all subunits of MSP1 are released except for MSP1-19 (Holder et al., 1984, Kauth et al., 2003). MSP1-19 contains two tandem EGF domains that are compact and rigid, and is linked to the membrane by a GPI anchor (Fig. 2C) (Chitarra et al., 1999, Morgan et al., 1999). The
structure of MSP1-19 in complex with an antibody that effectively coats the merozoite surface identified a region of MSP1-19 that is exposed on the merozoite (Pizarro et al., 2003). This region may contact Band 3, the erythrocyte receptor for MSP1 (Goel et al., 2003).

**The CSP surface coat on sporozoites**

CSP comprises the surface coat of the sporozoite during its journey through the mosquito salivary glands to invasion of human hepatocytes (Kappe et al., 2004). CSP possesses a distinct domain architecture: an N-terminal domain, a sequence designated Region I, a stretch of tetra-amino acid repeats, a C-terminus comprised of two Regions designated II and III, and an evolutionarily conserved thrombospondin type-I repeat (TSR) domain. A GPI anchor links CSP to the membrane (Wang et al., 2005).

The domains of CSP function in two distinct steps of the sporozoite’s life cycle. Initially, the CSP N-terminal domain/Region I binds heparan sulfate on the mosquito salivary glands (Sidjanski et al., 1997, Sinnis et al., 2007, Ghosh et al., 2009b, Armistead et al., 2011). Upon initiation of the mosquito blood meal, sporozoites are injected from the salivary glands into the human and migrate to the liver. Within the liver, the second step of CSP host cell recognition depends on proteolytic removal of the N-terminal domain and timed exposure of the C-terminal TSR domain (Coppi et al., 2005, Coppi et al., 2011). This exposure is correlated with recognition of increased sulfation of heparan sulfate proteoglycans on hepatocytes (Coppi et al., 2007, Coppi et al., 2011). Mutant sporozoites that constitutively express the cleaved version of CSP, exposing the TSR domain, continually migrate within the dermis and do not reach the liver. This suggests that the N-terminal domain prevents the CSP C-terminal region from binding improper receptors in the dermis, allowing the sporozoite to properly target the liver (Coppi et al., 2011). Retaining the N-terminus prior to hepatocyte recognition also shields the functional
C-terminal binding region from antibody recognition, representing an *in vivo* structural mechanism for protection of critical binding domains (Coppi *et al.*, 2011).

Ultrastructural analysis suggests that CSP maintains a flexible rod-like structure (Plassmeyer *et al.*, 2009). Small percentages of CSP appear to form dimers and oligomers in solution, suggesting that intermolecular contacts may function in sporozoite coat formation (Plassmeyer *et al.*, 2009). Extensive interdomain contacts are formed between the TSR domain and Region III (Fig. 2D) (Doud *et al.*, 2012). However, the TSR and Region III construct is monomeric in solution. Thus, CSP oligomerization may depend on the N-terminal domain or Region I and/or Region II, and internal repeat regions have been proposed to interact to form a protective coat for the parasite (Godson *et al.*, 1983).

**Surface proteins that adhere to cellular receptors during invasion**

After release from the liver into the bloodstream, the merozoite recognizes RBCs and must form a tight link with the host cell membrane. Two distinct protein families are involved at this stage: the Erythrocyte-binding-like (EBL) and the Reticulocyte-binding-like Protein Homologue (RH). Members of both families engage specific RBC receptors for invasion (Sim *et al.*, 1990, Adams *et al.*, 1992).

The EBL family contains a conserved domain architecture (Adams *et al.*, 1992). Receptor binding requires essential adhesive domains unique to Plasmodium termed Duffy Binding-like (DBL) found in Region II (RII) of EBL proteins. *P. falciparum* contains multiple EBL family members (PfEBA-175, PfEBA-140, PfEBL-1 and PfEBA-181), each containing two tandem DBL domains. In contrast, *P. vivax* is thought to be limited to a sole EBL member, Duffy Binding Protein (PvDBP), which contains a single DBL domain. However, sequencing of field isolates of *P. vivax* have revealed certain isolates carry a duplication of the PvDBP gene (Menard
et al., 2013), and others carry a novel EBL ligand that also has a single DBL domain (Hester et al., 2013). The DBL domains have a characteristic boomerang shape stabilized by extensive disulfide bridging (Tolia et al., 2005, Singh et al., 2006, Higgins, 2008, Khunrae et al., 2009, Batchelor et al., 2011, Juillerat et al., 2011, Lin et al., 2012, Vigan-Womas et al., 2012, Malpede et al., 2013). In addition to Region II, EBL proteins contain a segment of uncharacterized structure (Regions III-V), a structured C-terminal cysteine rich domain (Region VI), a transmembrane domain, and a cytoplasmic region (Adams et al., 1992, Withers-Martinez et al., 2008).

PfEBA-175-RII contains two tandem DBL domains that bind Glycophorin A in a sialic acid dependent manner (Camus et al., 1985, Sim et al., 1990, Klotz et al., 1992, Orlandi et al., 1992, Sim et al., 1994). The crystal structure of PfEBA-175-RII in complex with a sialic acid containing glycan revealed the receptor-binding pockets are formed at the dimer interface of PfEBA-175-RII (Fig. 2E) (Tolia et al., 2005). This suggests a dimer of PfEBA-175 assembles around dimeric Glycophorin A during invasion. Multimeric assembly of PfEBA-175 enhances binding to Glycophorin A and is augmented by regions outside of RII (Salinas et al., 2013, Wanaguru et al., 2013).

The P. vivax EBL member, PvDBP-RII, shares similar molecular structures and functional characteristics with PfEBA-175-RII (Tolia et al., 2005, Batchelor et al., 2011, Batchelor et al., 2013, Wanaguru et al., 2013). PvDBP-RII is monomeric in the absence of its receptor, the Duffy Antigen Receptor for Chemokines (DARC), and dimerization of PvDBP-RII is driven by receptor binding (Fig. 2F) (Batchelor et al., 2011). Crystal structures of PvDBP-RII in complex with the ectodomain of DARC, and complementary solution studies, demonstrated formation of two distinct complexes: a heterotrimer of two PvDBP-RII and one DARC, and a heterotetramer
of two PvDBP-RII and two DARC (Batchelor et al., 2013). These complexes are intermediates in a multi-step binding mechanism. In both structures, DARC is sandwiched by two PvDBP-RII molecules facilitating receptor-induced dimerization, suggesting a conserved binding mechanism between two EBL ligands despite distinct dimeric architectures (Figure 2E and 2F). However, the receptor binding sites and dimer interfaces lie on distinct faces of the DBL domains.

Crystal structures of PfEBA-140 Region II also identified receptor binding regions (Fig. 2G) (Lin et al., 2012, Malpede et al., 2013). PfEBA-140-RII is monomeric in the absence of its receptor Glycophorin C (Lobo et al., 2003, Lin et al., 2012, Malpede et al., 2013) and further studies of receptor-bound PfEBA-140 and other EBL ligands are necessary to determine if multimeric assembly occurs upon receptor binding. The receptor binding pockets in PfEBA-140 are located in a distinct region of the DBL fold compared to the pockets used by either PfEBA-175-RII or PvDBP-RII. This suggests that the DBL fold can create multiple binding pockets to bind a wide variety of receptors. PfEBA-140 also exhibits polymorphisms that affect receptor-specificity and/or binding affinity (Mayer et al., 2002, Maier et al., 2009), and a polymorphism that maps to one receptor-binding pocket suggests a structural basis for altered specificity (Malpede et al., 2013). Additional merozoite surface proteins containing the DBL domain bind RBCs and maintain the conserved DBL architecture (Hodder et al., 2012). The specific role of these DBL containing proteins during invasion is not completely understood.

EBL ligands function primarily in the blood stage, but one unique member, designated Apical membrane antigen/erythrocyte binding like (MAEGL), enables sporozoite entry into the mosquito salivary glands (Blair et al., 2002, Kariu et al., 2002, Fu et al., 2005). Sporozoites lacking MAEGL cannot attach to the salivary glands, but retain normal motility, supporting a specific role in attachment to host cells (Kariu et al., 2002). Unlike most EBL proteins that
contain DBL domains, the tandem adhesive domains in MAEBL contain homology to domains observed in Apical Membrane Antigen -1 (AMA-1), described below. MAEBL may be involved in a similar function at the moving junction as AMA-1, and may incorporate functional elements of AMA-1 and the EBL family.

Recent studies have illuminated the structure of the RH family of invasion ligands, and have focused on RH5 due to its proposed essentiality (Crosnier et al., 2011). Initial efforts illuminated a low resolution surface envelope of the RBC binding element from a RH family member (P. yoelli235) and suggested that this segment may resemble Region II of the EBL family (Gruber et al., 2011). However, the recent crystal structure of RH5 in complex with its receptor, Basigin, exhibited a novel invasion protein fold consisting helical bundles presenting receptor interaction interfaces near the N-terminus (Chen et al., 2014, Wright et al., 2014). Structures of RH5 in complex with inhibitory antibodies illuminated targeted epitopes localized to the receptor binding interface, similar to the results observed for the antibody response to EBL ligands (Batchelor et al., 2011, Batchelor et al., 2013, Chen et al., 2013).

**Adhesion during membrane invagination to form the parasitophorous vacuole**

In addition to forming a connection with the host cell membrane, the parasite must maintain this link during membrane invagination and engulfment into the parasitophorous vacuole, where it will develop internally. This membrane invagination requires the formation of the moving junction, a ring-shaped link between parasite and host membranes that begins at the apical end of the merozoite and migrates to the posterior end as the parasite invades. AMA-1 is a unique adhesin that is released during invasion, and is observed at the moving junction (Triglia et al., 2000, Lamarque et al., 2011). AMA-1 is expressed on sporozoites, and antibodies targeting this adhesin inhibit hepatocyte invasion, suggesting a functional role in the liver stage (Silvie et al.,
2004). Initial attempts to knockout AMA-1 were unsuccessful, however, more recent studies suggest AMA-1 is dispensable for blood stage growth (Bargieri et al., 2013).

AMA-1 is a type I integral membrane protein, and the extracellular portion is composed of three Domains (I, II, and III) defined by disulfide bridging (Hodder et al., 1996). The cytoplasmic region is conserved amongst Plasmodium species, and contains two C-terminal tyrosines that are proposed to function in signaling through phosphorylation, potentially providing communication with downstream effectors (Remarque et al., 2008). Domains I and II of AMA-1 adopt the evolutionarily conserved PAN domain fold, which mediates diverse protein-protein and protein-carbohydrate interactions (Fig. 2H) (Tordai et al., 1999, Bai et al., 2005, Pizarro et al., 2005). Domain III adopts a novel fold (Pizarro et al., 2005).

AMA-1 is not known to engage a host cell receptor. Instead, it binds the parasite expressed RON2, which is secreted from the merozoite rhoptries into the RBC membrane during invasion (Cao et al., 2009, Srinivasan et al., 2011). Crystallization with a peptide comprising the minimal RON2 binding domain clarified a specific hydrophobic groove of Domain I that engages the receptor (Vulliez-Le Normand et al., 2012). Displacement of a loop in Domain II is required for RON2 binding to AMA-1, and this loop may function to protect the binding site from host immunity until RON2 is available (Vulliez-Le Normand et al., 2012). AMA-1 binds the RON2 peptide with nanomolar affinity, and this interaction is strengthened by high local concentration and clustering of several independent AMA-1/RON2 interactions to facilitate moving junction formation (Vulliez-Le Normand et al., 2012). This clustering of interactions provides necessary avidity to maintain the parasite’s link to the host membrane as the parasite propels itself into a host cell.
The link between host cell receptors and the parasite's actin motor

The formation of the moving junction is accompanied by parasite movement that links an external host cellular receptor with the parasite's internal actin motor (Kappe et al., 1999, Baum et al., 2006). The thrombospondin related anonymous protein (TRAP) family provides this connection. TRAP family members contain a combination of adhesive evolutionarily conserved Von-Willebrand Factor Type A (VWA) and TSR domains. The cytoplasmic portion of TRAP members links to the actin motor through aldolase providing movement (Fig. 2I) (Buscaglia et al., 2003, Kappe et al., 2004, Bosch et al., 2007). Sporozoites deficient in the hallmark member of this family, TRAP, are capable of adhering to host cells, suggesting that TRAP function is required for invasion but not attachment (Sultan et al., 1997).

The VWA and TSR domains of PfTRAP and PvTRAP have been captured in various conformations by x-ray crystallography (Song et al., 2012). The VWA domains appear rigid but can adopt open and closed conformations with the closed conformation correlated with divalent cation binding. In contrast, the TSR domains are highly flexible, evidenced by complete disorder in PfTRAP and two distinct orientations in the open conformations of PvTRAP. This led the authors to suggest that TSR domains become structured upon receptor binding and the observed conformational changes may be necessary to initiate motility into the cell (Song et al., 2012).

An individual TRAP member is proposed to function at different stages of the parasite's life cycle (Baum et al., 2006). Movement through the mosquito midgut cellular barrier is the function of CTRP binding to laminin, as genetic deletion of CTRP severely disrupts this ability (Yuda et al., 1999a, Yuda et al., 1999b, Mahairaki et al., 2005). TRAP, present on the sporozoite, is required for movement into the mosquito salivary glands via binding to Saglin and also mediates hepatocyte invasion (Sultan et al., 1997, Wengelnik et al., 1999, Kappe et al.,
The parasite must also actively invade RBCs during blood stage growth, and MTRAP provides motor function at this juncture. The receptor for MTRAP has been defined as Semaphorin-7A, and the two appear to bind in a 2:2 complex (Bartholdson et al., 2012). Ultrastructural analysis demonstrates that MTRAP maintains an extended conformation that may provide flexibility that facilitates the active movement of the parasite into a host cell (Uchime et al., 2012). This flexible form may represent the overall shape of the TRAP family.

**Adhesive proteins exported to the infected RBC surface**

Within the RBC, the parasite actively exports variant surface antigens to the RBC membrane (Leech et al., 1984, Baruch et al., 1995, Smith et al., 1995, Su et al., 1995). The var gene family encodes erythrocyte membrane protein 1 (PfEMP1), which provides critical adhesive properties that provide protection from immune function (Baruch et al., 1995, Smith et al., 1995). The var gene family in *P. falciparum* contains approximately 60 unique variants in each haploid parasite genome, of which one variant is predominantly expressed at a given time (Chen et al., 1998, Scherf et al., 1998, Gardner et al., 2002). Each member contains an N-terminal domain along with a varied combination of two major Plasmodium specific adhesion domains: the DBL fold, and a cysteine rich interdomain region (CIDR) that is also highly alpha helical (Fig. 2J) (Smith et al., 2000, Heddini et al., 2001, Higgins, 2008, Klein et al., 2008, Khunrae et al., 2009). Similar to the DBL domains of merozoite adhesins, the PfEMP1 DBL domains appear to have a wide receptor range. The CIDR domains are also involved in receptor recognition, further increasing the receptor repertoire engaged by PfEMP-1.

Solution structural analysis by Small-angle X-ray Scattering (SAXS) of a recombinant full length extracellular region of one PfEMP1 variant, IT4VAR13, showed a flexible, extended molecule (Brown et al., 2013). In contrast, examination of a second PfEMP1 variant,
VAR2CSA, showed that the DBL-CIDR domains assemble into higher order structures that may involve domain-domain interactions (Srivastava et al., 2010). The full length extracellular region of VAR2CSA was required for highest affinity receptor binding, supporting a crucial role for interdomain interactions (Srivastava et al., 2010). Crystal structures of constructs containing the tandem N-terminus, DBL, and adjacent CIDR domain displayed extensive contacts between the individual elements of the PfEMP1 architecture (Juillerat et al., 2011, Vigan-Womas et al., 2012). The functional role for inter-domain contact is not clear, however, the domain-domain contacts represent a view that contrasts the previously invoked model of “beads on a string,” in which all PfEMP1 domains are flexible and capable of binding individual, unique receptors. Although individual domains are proposed to bind unique receptors, the role of each domain as a piece of the whole PfEMP1 protein is not fully understood.
1.3 Targeting adhesion proteins in vaccines

Exposure on the parasite surface marks *Plasmodium* adhesins as prime targets for host immunity and vaccines. The focus of current anti-malaria vaccines on full length adhesins or complete binding domains grants the immune system with access to decoy and non-inhibitory epitopes (Chen *et al.*, 2013), diminishing the production of inhibitory antibodies. Limiting or eliminating access to decoy and non-functional epitopes is critical towards developing rapid and effective immunity. Structural work has successfully defined conserved receptor binding sites and multimeric interfaces that can be specifically targeted to focus an antibody response. Functional interfaces have been defined by crystallizing the adhesin with receptors, and with characterized inhibitory antibodies. Techniques that focus the immune response to target specific epitopes are currently in development. Specific epitope targeting can be achieved by mutating immune-dominant non-inhibitory epitopes and by shrouding non-inhibitory epitopes with glycosylation (Ntumngia *et al.*, 2012, Sampath *et al.*, 2013). Similar approaches have effectively defined broadly-neutralizing epitopes to viral antigens, supporting the efficacy of these techniques (Corti *et al.*, 2013). Continued structural effort to identify critical and conserved contacts between parasite and host proteins provides an excellent opportunity in the development of vaccines that elicit strain-transcendent, highly inhibitory antibodies to malaria parasites.

**RBC invasion ligands as targets for vaccination**

The process of merozoite recognition and invasion of RBCs occurs in distinct stages (Cowman *et al.*, 2006, Cowman *et al.*, 2012). Initial low affinity recognition of the RBC is followed by reorientation of the merozoite with its apical tip facing the RBC membrane. This orientation initiates release of micronemes and rhoptries, organelles containing proteins that form an essential tight junction between parasite and host cell membranes and allow the
merozoite to insert itself into the cell, forming a parasitophorous vacuole for internal growth and
development (Aikawa et al., 1978, Bannister et al., 1990, Weiss et al., 2015). Targeting
interactions during recognition and insertion into the RBC provides an excellent intervention
strategy, as it prevents the parasite from expanding its population and limits the formation of
gametes, thus reducing or eliminating symptoms and transmission to new hosts (Sheehy et al.,
2013). The following thesis work examines the atomic and molecular basis of invasion for two
unique ligands of the EBL family, PfEBA-140 and PvDBP, focusing on defining receptor
interaction interfaces, determinants of receptor specificity, and biophysical binding mechanisms.
The ultimate objective of this work is to define functional elements of DBL domains for rational
targeting in a blood-stage malaria vaccine.
1.4 Figures

**Figure 1.** Domain architectures of adhesive proteins functioning at different parasite life stages. For families with a varied number and/or organization of adhesive domains, the most well characterized member of the family is shown. The domains are color coded and identified in the two boxes within the figure.
Figure 2. Crystal structures define adhesive folds used by the malaria parasite.

A. The structure of Pf12, representing the s48/45 domain. Members of the 6-cys family exhibit a range in their number of tandem s48/45 domains and function at multiple life cycle stages.

B. p25 utilizes four tandem EGF-like domains for adhesion (left). The four EGF domains are shown on the left in different colors for clarity. Extensive contact between p25 monomers was observed in the crystal packing arrangement, and these contacts are proposed to play a role in parasite surface coat formation (crystal packing arrangement shown on the right). The middle p25 monomer, shown in red, is equivalent to the p25 monomer shown on the left, while adjacent, contacting monomers are shown in black and grey.
C. MSP1-19 contains two tandem EGF-like domains, shown in red, involved in RBC binding. MSP1-19 displays extensive contact between the two EGFs, resulting in a rigid structure that contrasts other tandem EGF domain structures.

D. CSP Region III-TSR forms a rigid domain designated the α-TSR. Region III (grey) and the TSR (green) make extensive contacts.

E. PfEBA-175 engages its receptor Glycophorin A as a dimer. DBL domains are shown in blue for one PfEBA-175 monomer, and in grey for the second monomer that forms the dimeric complex during receptor engagement. The parasite membrane is shown in grey; the host RBC membrane is shown in red.

F. Binding of receptor DARC to PvDBP drives dimerization of this complex. The sole DBL domain of RII is shown in blue, while the contacting DBL domain from a second PvDBP is shown in grey. The parasite membrane is shown in grey; the host RBC membrane is shown in red.

G. PfEBA-140 appears to bind as a monomer to its receptor Glycophorin C. The tandem DBL domains of RII are shown in blue. The parasite membrane is shown in grey; the host RBC membrane is shown in red.

H. AMA-1 (orange/brown) binds the parasite expressed RON2 (purple), a member of the RON complex, which is released by the parasite into the RBC during invasion. AMA-1 is linked to cytoplasmic aldolase (light green) within the parasite. The parasite membrane is shown in grey; the host RBC membrane is shown in red.

I. The link to the parasite’s internal actin motor through cytoplasmic aldolase (light green) is formed by TRAP, with functions on the sporozoite. The VWA domain is shown in cyan and the TSR domain in green. A unique member of the TRAP family functions at each parasite life stage and utilizes a combination of the VWA and TSR domains. The parasite membrane is shown in grey; the host cell membrane (mosquito salivary gland and human hepatocyte), is shown in yellow.

J. During growth in the RBC, the parasite exports PfEMP1 to the RBC surface, where these proteins utilize a combination of the DBL (blue) and helical CIDR (brown) domains to adhere to a wide range of human surface receptors. The N-terminal element (purple) makes extensive contact with the DBL domain.
1.5 References


Coppi, A., Tewari, R., Bishop, J.R., Bennett, B.L., Lawrence, R., Esko, J.D., et al. (2007). Heparan sulfate proteoglycans provide a signal to Plasmodium sporozoites to stop migrating and productively invade host cells. Cell Host & Microbe 2, 316-327.


Chapter 2

Identification of a receptor binding interface and determinants of receptor specificity for PfEBA-140 RBC invasion
Preface

The following work was performed by me, Daniel H. Lin, Joseph D. Batchelor, and Niraj H. Tolia. D.H.L cloned, purified, and crystallized RII PfEBA-140, and collected, processed, and refined X-ray diffraction data. I analyzed the structure to identify a putative receptor binding interface and specific charged amino acids involved in the binding interaction. I performed the functional RBC binding experiments that illuminated critical receptor binding residues and completed the SAXS studies that identified PfEBA-140 as a monomer in solution. J.D.B and N.H.T analyzed SAXS data, refined X-ray data, and analyzed the structure. I wrote the majority of the manuscript and prepared the figures; all authors contributed to the paper writing, discussing results, and reviewing the manuscript. This chapter was published in its entirety: [Lin, D.H.*, Malpede, B.M.*, Batchelor, J.D., and Tolia, N.H. (2012) Crystal and Solution Structures of Plasmodium falciparum Erythrocyte-binding Antigen 140 Reveal Determinants of Receptor Specificity during Erythrocyte Invasion. Journal of Biological Chemistry 287, 36830]

*Authors contributed equally
Acknowledgements

We thank Nichole Salinas for assistance with experiments; Dr. Jay Nix of the Molecular Biology Consortium Beamline 4.2.2 of the Advanced Light Source (ALS) for help with data collection. The ALS is supported by the Director, Office of Science, Office of Basic Energy Sciences of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231; G. Hura, K. Dyer and the Structurally Integrated Biology for Life Sciences Team at the ALS for assistance in SAXS data collection funded by DOE IDAT grant contract number DE-AC02-05CH11231. This project was supported by grants from the National Institutes of Health (NIH) and the Edward Mallinckrodt, Jr. Foundation to Niraj H. Tolia, and is supported by a National Science Foundation Graduate Research Fellowship awarded to me under Grant No. DGE-1143954.
2.1 Abstract

Erythrocyte-binding antigen 140 (PfEBA-140) is a critical *Plasmodium falciparum* erythrocyte invasion ligand that engages Glycophorin C on host erythrocytes during malaria infection. The minimal receptor binding region of PfEBA-140 contains two conserved Duffy binding-like (DBL) domains, a fold unique to Plasmodia species. Here, we present the crystal structure of the receptor binding region of PfEBA-140 at 2.4 angstrom resolution. The two domain binding region is present as a monomer in the asymmetric unit, and the structure reveals novel features in PfEBA-140 that are likely determinants of receptor specificity. Analysis by small-angle X-ray scattering demonstrates that the minimal binding region is monomeric in solution, consistent with the crystal structure. Erythrocyte binding assays show that the full length binding region containing the tandem DBL domains is required for erythrocyte engagement, suggesting that both domains contain critical receptor contact sites. The electrostatic surface of PfEBA-140 elucidates a basic patch that constitutes a putative high affinity binding interface spanning both DBL domains. Mutation of residues within this interface results in severely diminished erythrocyte binding. This study provides insight into the structural basis and mechanism of PfEBA-140 receptor engagement and forms a basis for future studies of this critical interaction. In addition, the solution and crystal structures allow the first identification of likely determinants of erythrocyte receptor specificity for *P. falciparum* invasion ligands. A complete understanding of the PfEBA-140 erythrocyte invasion pathway will aid in the design of invasion inhibitory therapeutics and vaccines.
2.2 Introduction

Erythrocyte invasion by *Plasmodium* species is mediated by integral membrane proteins of the erythrocyte binding ligand (EBL) family. During invasion, EBL proteins bind irreversibly and specifically to erythrocyte receptors to create a tight junction between host and parasite membranes. This interaction facilitates merozoite entry into the red blood cell. *Plasmodium falciparum* has a sophisticated invasion machinery with several EBL proteins that each bind a different erythrocyte receptor in a sialic acid dependent manner (Cowman *et al.*, 2006). Erythrocyte-binding antigen 175 (PfEBA-175), erythrocyte-binding ligand 1 (PfEBL-1), and erythrocyte-binding antigen 140 (PfEBA-140) bind Glycophorin A, B, and C, respectively (Orlandi *et al.*, 1992, Lobo *et al.*, 2003, Mayer *et al.*, 2009). A fourth member of this family, erythrocyte-binding antigen 181 (PfEBA-181), binds an unknown receptor (Gilberger *et al.*, 2003). The EBL family members are composed of two cysteine rich regions designated region II (RII) and region VI and contain a type I transmembrane domain and short cytoplasmic domain (Adams *et al.*, 1992). Receptor binding has been localized to RII for all members. In the *P. falciparum* EBL family, this region contains two tandem Duffy binding-like (DBL) domains, F1 and F2. The DBL protein fold is unique to *Plasmodia*, and is able to recognize and tightly bind a diverse array of host-cell receptors. In addition to their critical role during invasion, DBL domains also mediate microvasculature adherence of infected erythrocytes by erythrocyte membrane protein 1 (PfEMP1), a phenomenon directly associated with severe malaria (Kraemer *et al.*, 2006).

It is unknown how the EBL proteins utilize such a highly conserved domain structure to recognize different erythrocyte receptors and thus provide *P. falciparum* with multiple pathways for invasion. In addition, the role of each individual erythrocyte invasion pathway during natural
infection is not fully understood. However, it has been observed that Gerbich negativity is present at high frequency in regions of Papau New Guinea where infection with *P. falciparum* malaria is common (Maier *et al.*, 2003). Gerbic h negative individuals have a deletion of exon 3 in the GPC gene that prevents PfEBA-140 erythrocyte binding and invasion. This observation provides strong evidence that severe malaria has selected for this mutation and illustrates the significance of erythrocyte invasion mediated by PfEBA-140.

As the tandem DBL domains in RII of the four *P. falciparum* EBL proteins can independently bind erythrocytes, they are the focus of combinatorial vaccine efforts. To understand how *P. falciparum* uses the DBL protein fold to recognize different erythrocyte receptors during invasion, we determined the crystal structure and examined the erythrocyte binding profile of RII PfEBA-140. In addition, the solution structure and oligomeric state of this invasion ligand were examined using small-angle X-ray scattering (SAXS). The results presented here elucidate likely determinants of receptor specificity within the EBL family and provide insight into the structural basis of erythrocyte binding by the critical invasion ligand PfEBA-140.
2.3 Experimental Procedures

Protein expression and purification

A codon-optimized construct containing amino acids 143-740 of RII PfEBA-140 with three point mutations (Ser303Ala, Thr469Ala, Ser727Ala) was cloned for expression. The three mutations were introduced to avoid aberrant glycosylation at putative N-glycosylation sites during expression in mammalian cells that could otherwise impact protein homogeneity. This construct provided high yields of pure protein from *Escherichia coli* and was thus used for crystallization. These amino acid changes do not affect protein function, as demonstrated by erythrocyte binding assays described below. The construct was expressed as inclusion bodies in *E. coli* and recovered using 6M guanidinium hydrochloride. Following recovery, 100mg/L of denatured protein was rapidly diluted in 50 mM Tris pH 8.0, 10 mM EDTA, 200 mM arginine, 0.1 mM PMSF, 2 mM reduced glutathione, and 0.2 mM oxidized glutathione. After 48 hours of refolding at 4°C, RII PfEBA-140 was concentrated using Amicon centrifugal filters and purified by size exclusion and ion-exchange chromatography.

Crystallization, data collection, and structure determination

Crystals were grown using the hanging drop vapor diffusion method by mixing 1uL protein at 7.5mg/mL with 1uL of reservoir containing 20% PEG 8000 and 0.1 M HEPES pH 7.5. Initial crystal hits observed in precipitant screens (Qiagen) were used as seeds to optimize crystal growth. Seeds were generated by transferring the entire crystal drop into 10 μl 20% PEG 8000 and 0.1 M HEPES pH 7.5 and vortexing the sample. Crystals were sent for remote data collection at beamline 4.2.2 at the Advanced Light Source, Lawrence Berkeley National Laboratory. Cyroprotectant composed of 30% glycerol, 17.5% PEG 8000, and 0.1 M HEPES pH 7.5 was introduced gradually into the drop by pipetting. Crystals were isolated with nylon loops.
and stream frozen prior to transport.

The RII PfEBA-140 structure was solved using molecular replacement using RII PfEBA-175 as a model in BALBES (Long et al., 2008). Automated refinement was performed with PHENIX (Adams et al., 2002) and manual model building was performed with Coot (Emsley et al., 2004). Refinement was completed once low R-factors and good geometry were obtained (Table 1). The structure described has been deposited in the Protein Data Bank with accession code 4GF2.

**Small-angle X-ray Scattering (SAXS)**

SAXS data was collected at the SIBYLS beamline 12.3.1 at the ALS using standard procedures. 30 μl of RII PfEBA-140 at 1 mg ml\(^{-1}\) was automatically loaded into the cuvette with a Hamilton Syringe robot (Hura et al., 2009). Radiation damage was assessed by overlaying short exposures using PRIMUS (Konarev et al., 2003). CRYSOL (Svergun et al., 1995) was used to compare experimental profiles to the crystal structure and DAMMIF was used for *ab initio* model generation (Franke et al., 2009). Structures and SAXS reconstructions were aligned using SUPCOMB20 (Kozin et al., 2001).

**Erythrocyte Binding Assay**

RII PfEBA-140 containing the three mutations (Ser303Ala, Thr469Ala, Ser727Ala) was fused to a C-Terminal GFP and cloned into pRE4 for surface expression on HEK-293T cells. Monolayers of HEK-293T cells grown in 3.5cm wells were transfected with 2.7ug of plasmid DNA in polyethyleneimine. The erythrocyte binding assay was performed 20 hours after transfection. Transfected cells were incubated with normal human erythrocytes at 2% hematocrit for 2 hours. Following incubation, the cells were washed 3 times with Phosphate Buffered Saline (PBS). In each experiment, 3 individual wells of HEK-293T cells were transfected with each construct. Binding phenotypes were assessed for 10 fields of view from each of the 3 wells of
transfected cells (for a total of 30 images for each construct).
2.4 Results

Overall structure of PfEBA-140 Region II

To define the structural basis of PfEBA-140 receptor specificity, we solved the crystal structure of Region II PfEBA-140 (Fig. 1A and Table 1). For simplicity, Region II PfEBA-140 will be referred as RII PfEBA-140. RII PfEBA-140 is present as a monomer in the asymmetric unit, and no oligomeric contacts with potential physiological relevance are observed (Fig. 1A). The two DBL domains, F1 (residues 143-422) and F2 (residues 447-740), are connected by a short, helical linker. Each DBL domain is composed of three subdomains (Fig. 1B), and contains a unique disulfide bridging pattern relative to other DBL family members (Fig. 1C). The F1 and F2 domains of PfEBA-140 are structurally similar to other DBL domains of *Plasmodium* invasion ligands (Fig. 1D).

**PfEBA-140 Erythrocyte Binding Requires Both DBL Domains**

The presence of two structurally conserved domains within the minimal binding region suggests that each domain may be able to independently engage erythrocytes. To assess the function of the individual DBL domains, full length RII PfEBA-140 and constructs containing each individual DBL domain were tested for erythrocyte binding by rosetting assay. Full length RII PfEBA-140, corresponding to the construct used for crystallization, binds extensively to erythrocytes (Fig. 2A). In contrast, neither DBL domain is capable of independently engaging erythrocytes (Fig. 2B), suggesting that both domains make essential contacts with GPC during invasion. This result is supported by examination of the electrostatic surface of RII PfEBA-140, which elucidates a putative high affinity binding interface (Fig. 2C). This region forms an arch of overall positive charge that spans the two DBL domains and would provide an ideal interaction surface for engagement of the highly glycosylated, acidic GPC.
To assess the functional importance of the putative binding interface, individual amino acid residues within the basic patch were mutated to alanine and tested for deficient erythrocyte binding. Large polar and/or charged residues are often involved in glycoprotein interactions, thus we focused on these residues for testing (Fig. 2C). Four residues were identified that when mutated to alanine resulted in greatly diminished or null erythrocyte binding (Fig. 2D). This result confirms the vital role of individual residues within the putative interface during receptor binding. To test the alternate side of the protein, residues on the face opposite of the basic patch were also mutated to alanine. Mutation of these residues had no effect on binding, supporting the specific functional role of residues within the basic interface (Fig. 2E).

*The structural basis of PfEBA-140 receptor specificity*

Unique structural elements distinguish RII PfEBA-140 from other DBL domain containing EBL ligands and are the likely basis of receptor specificity. All 26 cysteines in RII PfEBA-140 are involved in disulfide bonds, and two of the disulfide linkages are distinct from other characterized DBL domains (Fig. 1C). The modified disulfide pattern includes a linkage between cysteine 7 and cysteine 8, as well as a linkage between cysteine 9 and cysteine 12. In contrast, linkages between cysteine 7 and cysteine 9 as well as cysteine 8 and cysteine 12 are observed in all other characterized EBL members.

In addition to an altered disulfide pattern, several elements of the RII PfEBA-140 subdomain structure are unique among DBL domains. In RII PfEBA-175, β-fingers present in both DBL domains contain long loops that serve as critical dimeric contacts required for receptor binding (Tolia et al., 2005). In contrast, *P. vivax* Duffy-binding protein (PvDBP) and *P. knowlesi* Duffy-binding protein (PkDBP), which contain only a single DBL domain in their Region II, possess short loops within the β-finger region that do not seem to be important for the function of these
ligands (Singh et al., 2006, Batchelor et al., 2011). The β-finger of F1 PfEBA-140 is similar in length to the β-fingers of the PfEBA-175 F1 and F2 domains. In addition, the F1 β-fingers of PfEBA-140 and PfEBA-175 overlay quite well (Fig. 3A). However, the canonical β-finger motif in F1 PfEBA-140 is replaced with a novel alpha helical segment in the F2 domain (Fig. 3B).

A third critical difference is the orientation of subdomain 3 in F1 PfEBA-140 with respect to the rest of the DBL domain. In most DBL domains, subdomain 3 contains a kink in a long helix facilitated by a glycine residue (Gly185 in F1 PfEBA-175, Gly490 in F2 PfEBA-175, Gly397 in PvDBP, Gly394 in PkDBP-α and Gly627 in F2 PfEBA-140). This glycine is absent in F1 PfEBA-140 and without the inherent flexibility of this residue, the kink in subdomain 3 cannot form. The absence of this kink leads to a drastic change within the subdomain that propagates through RII PfEBA-140, resulting in a large difference in the hinge angle between the two DBL domains compared to RII PfEBA-175 (Fig. 3A and Fig. 3C). The splayed out DBL domains of RII PfEBA-140 would require a large hinge movement to create the mode of dimerization seen in the crystal structure of RII PfEBA-175 (Fig. 3A, inset).

**Solution structure of PfEBA-140**

The structure and oligomeric state of parasite ligands in solution is an important determinant of function. Both the related ligands PvDBP and PfEBA-175 appear to bind receptors as dimers (17, 18). We determined the solution structure and oligomeric state of RII PfEBA-140 by SAXS. The theoretical scatter for the monomeric PfEBA-140 crystal structure resulted in an excellent fit to the experimental SAXS profile with a $\chi^2$ of 1.46 (Fig. 4). Furthermore, an averaged *ab initio* reconstruction revealed a molecular envelope that closely resembles the monomeric structure. This result is consistent with the observed crystal contacts and hinge angle between the F1 and F2 domains suggesting RII PfEBA-140 is monomeric in solution (Fig. 4). Only monomeric
forms of PfiEBA-140 have been described here and there is no evidence for higher order oligomeric states in the absence of GPC.
2.5 Discussion

*P. falciparum* field isolates from endemic malaria regions actively utilize multiple erythrocyte invasion ligands, exemplifying the need to characterize each individual pathway (Lobo et al., 2004). Antibodies targeting PfEBA-140 are capable of inhibiting invasion, supporting the critical role of this ligand during invasion (Lopaticki et al., 2011). Furthermore, PfEBA-140 is naturally immunogenic. Serum isolated from patients infected with *P. falciparum* malaria is reactive to PfEBA-140, and RII was found to be the most immunoreactive element (Ford et al., 2007). The active immune response to PfEBA-140 provides strong support for its validity as a component of a combinatorial vaccine targeting invasion ligands. The structure described here forms a framework for future studies of the critical interaction between PfEBA-140 and erythrocytes. In addition, identifying the structural basis of receptor specificity within the EBL family is essential to characterizing the full range of invasion pathways utilized by *P. falciparum*. These results will thus aid in the rational development of invasion inhibitory therapeutics and vaccines.

The structural differences identified in RII PfEBA-140 suggest an altered mechanism of binding relative to PfEBA-175, and likely other EBL members. It has been proposed that dimerization is an important mechanistic component of EBL mediated invasion. Support for this proposed mechanism comes from the fact that PvDBP is monomeric in the absence of receptor and dimerizes upon binding its receptor, DARC (Batchelor et al., 2011). In addition, PfEBA-175 crystallized as a dimer and may engage GPA in the dimeric form observed in the crystal structure (Tolia et al., 2005). It is possible that PfEBA-140 follows this EBL pattern and engages GPC as a dimer while existing as a monomer in the absence of receptor. However, to bind GPC using the dimer architecture observed for PfEBA-175, RII PfEBA-140 would need to undergo a large
structural hinge movement and display dimeric contacts not observed in PfeBA-175. Specifically, the β finger motifs in each DBL domain of PfeBA-175 make dimeric contacts. The presence of a unique helical element in place of the canonical β finger motif in F2 PfeBA-140 may alter the dimeric contacts observed in PfeBA-175, resulting in a novel dimer conformation of PfeBA-140. It is equally possible that PfeBA-140 engages GPC as a monomer and that oligomeric state is an important determinant of receptor specificity.

Receptor recognition by PfeBA-140 is poorly understood; however, it is known that PfeBA-140 erythrocyte engagement is dependent on GPC glycans as well as the protein backbone. The proposed binding interface identified in the crystal structure spans both DBL domains (Fig. 2C). We demonstrated that both domains are required for erythrocyte binding and that mutating individual residues in each domain severely diminishes erythrocyte binding. These results suggest that each DBL domain forms essential contacts with GPC during invasion. In addition, the identification of the positive interface forms a basis for identifying the GPC binding site. It is probable that vital receptor binding interactions are located in this putative binding interface due to the concentrated presence of basic residues. The only known essential binding component on GPC is a solitary N-linked glycan at residue 8 (Mayer et al., 2006). Putative O-linked glycans are abundant on GPC, but their role in binding has not been examined in detail. The critical N-linked glycan may make contacts with both DBL domains, explaining the requirement for full length RII during binding. Alternatively, the N-linked glycan may bind with high specificity to one domain, while O-linked glycans and the protein backbone form essential contacts with the other domain. Further studies examining the interaction of RII PfeBA-140 with GPC are required to fully understand the structural and mechanistic basis of receptor recognition during invasion. In conclusion, we have provided insight into the PfeBA-140 binding mechanism,
elucidated an important receptor binding interface, and identified unique structural motifs in RII PfEBA-140 that form the basis of receptor specificity within the EBL family that allows *P. falciparum* to engage multiple host receptors during invasion.
2.6 Figures and Tables

Figure 1. Structure of RII PfEBA-140. A, Overall structure of RII PfEBA-140 as well as the location of individual subdomains within each DBL domain. F1 subdomain 1 is shown in bronze, subdomain 2 in orange, and subdomain 3 in dark orange. F2 subdomain 1 is shown in dark blue, subdomain 2 in blue, and subdomain 3 in light blue. The short helical linker between the two domains is shown in grey. B, Structures of the individual subdomains in RII PfEBA-140. Coloring is the same as in A. C, Representation of the disulfide bridging pattern of characterized EBL ligands displaying the two altered disulfides in RII PfEBA-140. D, Overlay of *Plasmodium* invasion ligand DBL domains with each DBL domain of RII PfEBA-140. The top layer displays F1 PfEBA-140 in orange; the bottom displays F2 PfEBA-140 in blue. F1 PfEBA-175 is shown in dark green, F2 PfEBA-175 in purple, PvDBP in light green, and PkDBP in brown.
Figure 2. PfEBA-140 erythrocyte binding requires both DBL domains. A, RII PfEBA-140 was expressed on the surface of HEK-293 mammalian cells and tested for erythrocyte binding. GFP was utilized to assess proper surface expression. A construct expressing only GFP does not bind erythrocytes (top panel). The full length RII construct, identical to the construct used for crystallization, is capable of extensively engaging erythrocytes (bottom panel, bound erythrocytes appear black around the mammalian cells). B, The individual DBL domains of PfEBA-140 were also tested for erythrocyte binding. The F1 (top panel) and F2 (bottom panel) DBL domains cannot independently engage erythrocytes. C, The electrostatic surface of PfEBA-140 RII elucidates a putative high affinity binding interface. The surface of the protein shown in the top panel has no clear region of concentrated charged residues (blue: positive charge, red: negative charge). However, when rotated 180 degrees, a region of positive charge that forms an arch between the two DBL domains is illuminated (middle panel). Surface charge potential is colored from +3.5eV to -3.5eV. The bottom panel displays the face of the protein containing basic patch identified as a putative binding interface. Residues mutated to alanine and tested for erythrocyte binding are identified with arrows and shown in black. D, Alanine mutations within the basic patch disrupt erythrocyte binding. E, In contrast to residues within the basic patch, alanine mutations on the opposite face of RII PfEBA-140 do not disrupt erythrocyte binding.
Figure 3. Structural determinants of receptor specificity within the EBL family. A, Overlay of PfEBA-140 (F1 domain in orange, F2 domain in blue) with PfEBA-175 (F1 domain in green, F2 domain in purple). The conserved F1 beta sheet motifs are outlined in black. The inset outlined in black displays two monomers of PfEBA-140 overlayed onto the PfEBA-175 dimer observed in the crystal structure (17). B, Overlay of the PfEBA-140 F1 and F2 DBL domains (left). The altered helical element in F2 is outlined in black. A close up view of this F2 helical element, overlayed onto the F1 beta sheet, is shown on the right. C, View of the altered helical orientation of F1 RII PfEBA-140 subdomain 3 (orange). This orientation occurs due to the absence of a glycine, which is conserved in other EBL members, at residue 325 in PfEBA-140. The kink observed in F1 PfEBA-175 (green) is shown for comparison.
Figure 4. Oligomeric state of RII PfEBA-140. RII PfEBA-140 is monomeric in solution in the absence of receptor. Shown here are the experimental and theoretical SAXS plots of scattering intensity versus scattering momentum. The inset displays an ab initio construct of PfEBA-140 (light blue) overlayed onto the crystal structure (F1 domain in orange, F2 domain in blue).
Table 1: Data collection and refinement statistics

<table>
<thead>
<tr>
<th></th>
<th>RII PfEBA-140</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data collection</td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>65.42, 76.48, 82.34</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90.00, 96.80, 90.00</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>20-2.4 (2.49-2.40)</td>
</tr>
<tr>
<td>R_{merge}</td>
<td>5.8% (78.9%)</td>
</tr>
<tr>
<td>I/σI</td>
<td>19.06 (2.07)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.4% (98.0%)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.72 (3.67)</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>20.0-2.4</td>
</tr>
<tr>
<td>No. reflections</td>
<td>31,507</td>
</tr>
<tr>
<td>R_{work} / R_{free} (%)</td>
<td>20.03 / 23.69</td>
</tr>
<tr>
<td>No. atoms</td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>5,339</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>90</td>
</tr>
<tr>
<td>Water</td>
<td>143</td>
</tr>
<tr>
<td>B-factors</td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>67.57</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>78.16</td>
</tr>
<tr>
<td>Water</td>
<td>53.73</td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.003</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>0.607</td>
</tr>
<tr>
<td>Ramachandran</td>
<td></td>
</tr>
<tr>
<td>Favored</td>
<td>97.46%</td>
</tr>
<tr>
<td>Allowed</td>
<td>2.54%</td>
</tr>
<tr>
<td>Disallowed</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

Values in parentheses are for highest-resolution shell.
2.7 References


Chapter 3

The molecular basis for sialic acid dependent receptor recognition by PfEBA-140
Preface

The following work was performed by me, Daniel H. Lin, and Niraj H. Tolia. D.H.L cloned, purified, and crystallized RII PfEBA-140 in complex with sialyllactose, and in conjunction with N.H.T. collected, processed, and refined x-ray diffraction data. I analyzed the structure for structural determinants of receptor specificity in the \textit{P. falciparum} EBL family and for comparison to other sialic acid binding proteins functioning in other parasites and the human immune system. I identified residues that interact with or form a structural base for sialic acid binding, and performed the functional RBC binding studies that defined the sialic acid binding pockets of the DBL domains. I prepared the figures and wrote the manuscript. This work was published in its entirety: [Malpede, B.M.*, Lin, D.H.*, and Tolia, N.H. (2013) Molecular Basis for Sialic Acid-dependent Receptor Recognition by Plasmodium falciparum Erythrocyte Binding Antigen 140/BAEBL. \textit{Journal of Biological Chemistry} 288, 12406-12415

*Authors contributed equally
Acknowledgements

We thank Nichole Salinas for assistance with experiments, Wandy Beatty for help with confocal imaging, David Sibley for constructive comments on the manuscript, the Structural Biology Center (SBC) at Argonne National Laboratory for use of beamline 19-ID at the Advanced Photon Source and Stephen L. Ginell and SBC beamline scientists for assistance during data collection. Argonne is operated by UChicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357. This project is supported by grant number AI080792 from the National Institutes of Health (NIH) and the Edward Mallinckrodt, Jr. Foundation to Niraj H. Tolia, and is supported by a National Science Foundation Graduate Research Fellowship awarded to me under Grant No. DGE-1143954.
3.1 Abstract

*Plasmodium falciparum* erythrocyte invasion is dependent on high-affinity recognition of sialic acid on cell-surface receptors. The Erythrocyte Binding-like (EBL) family of invasion ligands mediates recognition of sialic acid on erythrocyte glycoproteins. Erythrocyte Binding Antigen-140 (PfEBA-140/BAEBL) is a critical EBL ligand that binds sialic acid on its receptor Glycophorin C. We present here the crystal structure of the two domain receptor-binding region of PfEBA-140 in complex with a glycan containing sialic acid. The structure identifies two glycan-binding pockets unique to PfEBA-140 and not shared by other EBL ligands. Specific molecular interactions that enable receptor engagement are identified and reveal that the glycan-binding mode is distinct from that of Apicomplexan and viral cell surface recognition ligands as well as host immune factors that bind sialic acid. Erythrocyte binding experiments elucidated essential glycan contact residues and identified divergent functional roles for each receptor binding site. One of four polymorphisms proposed to affect receptor-binding was localized to a glycan binding site, providing a structural basis for altered erythrocyte engagement. The studies described here provide the first full description of sialic acid-dependent molecular interactions at the *P. falciparum* erythrocyte invasion interface and define a framework for development of PfEBA-140 based therapeutics, vaccines, and diagnostics assessing vaccine efficacy and natural immunity to infection.
3.2 Introduction

*Plasmodium falciparum* erythrocyte invasion requires targeted recognition of cell surface receptors by merozoite ligands (Camus *et al.*, 1985, Sim *et al.*, 1990). High affinity binding to erythrocyte receptors is critical to the invasion process and is mediated by the Erythrocyte Binding-like (EBL) family of proteins (Adams *et al.*, 1992). In *P. falciparum* the EBL family is composed of four known functional membrane embedded ligands that each target a specific erythrocyte receptor in a sialic acid dependent manner (Adams *et al.*, 1992). Receptor engagement by the merozoite allows the formation of an irreversible tight junction with the host membrane and subsequent formation of the parasitophorous vacuole (Dvorak *et al.*, 1975, Aikawa *et al.*, 1978). Binding interactions mediated by the EBL family are vital to the survival of the parasite and are thus targets of therapeutics and vaccines (Aravind *et al.*, 2003, Cowman *et al.*, 2006).

Each member of the *P. falciparum* EBL family contains an extracellular cysteine rich region composed of two Duffy binding-like (DBL) domains designated F1 and F2 (Mayor *et al.*, 2005). These two domains comprise the minimal binding region of the *P. falciparum* EBL ligands, and have been designated Region II (RII). *P. falciparum* Erythrocyte Binding Antigen 140 (PfEBA-140/BAEBL) is a member of the EBL family that binds erythrocytes via its cell surface receptor Glycophorin C (GPC) (Lobo *et al.*, 2003). Studies examining the erythrocyte binding capability of the individual RII PfEBA-140 DBL domains (F1 and F2) demonstrated that neither domain is sufficient to engage erythrocytes (Mayer *et al.*, 2002, Lin *et al.*, 2012). In addition, single amino acid mutations in either DBL domain severely disrupt erythrocyte binding, demonstrating an essential role for both DBL domains during invasion (Lin *et al.*, 2012).

Receptor recognition by *P. falciparum* invasion ligands is also characterized by the
sensitivity of the binding interaction to specific enzyme treatments of red blood cells (Miller et al., 1977, Mayer et al., 2001). Reports examining the binding profile of PfEBA-140 to enzyme treated erythrocytes have provided a basic understanding of how this ligand engages its receptor. PfEBA-140 receptor binding is trypsin and neuraminidase sensitive, but chymotrypsin resistant (Mayer et al., 2001). In addition, soluble sialic acid is not capable of inhibiting PfEBA-140 erythrocyte binding (Mayer et al., 2001). These results demonstrated that receptor glycans containing sialic acid are essential for receptor recognition, and that the protein backbone of GPC also plays a role in binding (Mayer et al., 2001, Narum et al., 2002, Jiang et al., 2009). GPC possesses several putative O-linked glycans and one known N-linked glycan, all containing the sialic acid sugar moiety required for PfEBA-140 binding. The solitary N-linked glycan is essential for GPC engagement, but the role of individual O-linked glycans is not clear (Mayer et al., 2006).

Numerous and widespread polymorphisms have been identified in Plasmodium invasion ligands, specifically PfEBA-175 and P. vivax Duffy Binding Protein (PvDBP) (Xainli et al., 2000, Baum et al., 2003, Gosi et al., 2008). In contrast, only four polymorphic residues have been identified in RII PfEBA-140, all of which are present in the F1 domain: Ile185Val, Asn239Ser, Lys261Arg/Thr, and Lys285Glu (Mayer et al., 2002). Polymorphic residue changes in PfEBA-140 modify this ligand’s binding profile to enzyme-treated erythrocytes. The altered binding profile suggests that PfEBA-140 is capable of interacting with other erythrocyte cell surface molecules and may mediate an invasion process independent of GPC (Thompson et al., 2001, Mayer et al., 2002). In addition to altering the binding profile, it has been shown that polymorphisms reduce the affinity of PfEBA-140 for erythrocytes (Maier et al., 2009). Recent evidence suggests that binding to glycosaminoglycans (GAGs) on the erythrocyte surface
promotes merozoite invasion (Boyle et al., 2010). PfEBA-140 is capable of binding heparin, and it has been proposed that cell surface GAGs may function as a secondary interaction element during PfEBA-140 mediated invasion (Kobayashi et al., 2010).

The role of individual EBL invasion ligands during in vivo erythrocyte invasion and blood stage growth is not fully understood. However, several studies have demonstrated the importance of PfEBA-140 during natural infection as both an invasion ligand and antigen. Individuals in malaria endemic regions mount a strong antibody response to PfEBA-140, and RII was the most immunoreactive element (Ford et al., 2007). This result provides evidence of PfEBA-140 expression and immune recognition in natural infections and supports the significance of this ligand as a member of a combinatorial vaccine. In addition to natural immunogenicity, antibodies targeting RII PfEBA-140 are capable of inhibiting invasion, which is strongly supportive of the functional role and antigenic properties of PfEBA-140 (Maier et al., 2003, Lopaticki et al., 2011). Furthermore, Gerbich negativity is observed at high frequencies in regions of endemic malaria in Papau New Guinea (Maier et al., 2003). This phenotype results from the loss of exon 3 within the GPC gene. The absence of this portion of GPC prevents PfEBA-140 from engaging erythrocytes and thus inhibits the invasion process. The prevalence of Gerbich negativity provides evidence that severe malaria has selected for this mutation and supports the importance of PfEBA-140 during natural infection (Maier et al., 2003).

To identify receptor binding sites in PfEBA-140 and define molecular interactions at the merozoite-erythrocyte invasion interface, we solved the crystal structure of RII PfEBA-140 in complex with a glycan containing sialic acid, the essential sugar component of GPC that is recognized during erythrocyte engagement. Two glycan binding pockets are identified, one in each DBL domain. Mutation of individual sialic acid contact residues disrupts erythrocyte
binding by RII PfEBA-140, confirming the functional role of the sugar binding pockets. Stark differences in the mutant erythrocyte binding phenotypes for the F1 and F2 domains suggest that each DBL domain performs a distinct function during erythrocyte engagement. One of the four polymorphic residues that affect PfEBA-140 receptor-binding is found in the base of the F1 glycan binding pocket. The localization of this residue provides insight into the structural basis of altered receptor binding observed for polymorphic variants. The structure also allowed for mapping of putative sulfate and GAG binding sites, which may represent true interaction elements that promote red blood cell binding and invasion. This study provides the first complete molecular and structural description of sialic-acid dependent interactions critical to the formation of the tight junction at the merozoite-erythrocyte invasion interface. Our results will thus aid in the design of novel therapeutics and diagnostics targeting this essential step in the parasite’s life cycle.
3.3 Experimental Procedures

Protein expression and purification

RII PfEBA-140 was expressed and purified as previously described (Lin et al., 2012). Briefly, RII PfEBA-140 was expressed in Escherichia coli and recovered from inclusion bodies using 6M guanidinium hydrochloride. After overnight denaturing, 100mg/L of denatured protein was rapidly diluted in 50 mM Tris pH 8, 10 mM EDTA, 200 mM arginine, 0.1 mM PMSF, 2 mM reduced glutathione, and 0.2 mM oxidized glutathione. The protein was refolded for 48 hours at 4°C. After refolding, RII PfEBA-140 was purified by ion-exchange and size exclusion chromatography and concentrated using Amicon centrifugal filters for crystallization.

Crystallization, data collection, and structure determination

Crystals of RII PfEBA-140 were grown using the hanging drop vapor diffusion method. The drops were produced by mixing 1uL of protein at 7.5mg/mL with 1uL of reservoir containing 20% PEG 8000 and 0.1 M HEPES pH 7.5. Crystals of RII PfEBA-140 alone were soaked with 10mM sialyllactose for 30 minutes to obtain the complex. Crystals were sent for remote data collection in cryoprotectant composed of 30% glycerol, 17.5% PEG 8000, and 0.1 M HEPES pH 7.5. The cryoprotectant was introduced gradually into the drop by pipetting and crystals were isolated with nylon loops and stream frozen prior to transport. Data was collected at beamline 19-ID at the Advanced Photon Source, Argonne National Laboratory and processed with XDS (Kabsch, 2010). The structure of RII PfEBA-140 bound to sialyllactose was solved using the unbound PfEBA-140 structure PDB 4GF2 as a model in Phaser (McCoy et al., 2007, Lin et al., 2012). Refinement and model building were performed with PHENIX (Adams et al., 2002) and Coot (Emsley et al., 2004) and refinement was completed once low R-factors and good geometry were obtained (Table 1). The structure described has been deposited in the Protein Data Bank.
Functional studies

RII PfEBA-140 was fused to a C-Terminal GFP in plasmid pRE4 for surface expression on HEK-293T cells. Individual alanine point mutants were produced using the QuikChange method and verified by plasmid sequencing. Monolayers of HEK-293T cells grown in 3.5cm wells were transfected with 2.7ug of plasmid DNA in polyethyleneimine. Erythrocyte binding of individual constructs was assayed 20 hours after transfection.

To assess the binding phenotypes of RII PfEBA-140 mutant constructs, transfected HEK-293 cells were incubated with normal human erythrocytes. For enzyme treatment experiments, normal erythrocytes at 50% hematocrit were treated with 0.1mg/mL trypsin, 0.1mg/mL chymotrypsin, or 5mU of *Vibrio cholera* neuraminidase at 37ºC for 2 hours and washed prior to use. Erythrocytes at 2% hematocrit were then overlayed onto the transfected HEK-293 cells for 2 hours. Following incubation with untreated or enzyme treated erythrocytes, the transfected HEK-293 cells were washed 3 times with Phosphate Buffered Saline (PBS) to remove unbound erythrocytes. For the mutation studies, three individual wells of HEK-293 cells were transfected with each mutant construct. Binding percentage was calculated as the number of rosette positive cells over the number of GFP positive cells and was normalized to wild type binding. Binding phenotypes were quantified over 10 fields of view from each of the 3 wells of transfected cells (a total of 30 images for each construct). Images were obtained using a Zeiss LSM 510 META Laser Scanning Microscope with an LD Achroplan 20X Korr DIC objective. Images were randomized prior to counting with ImageJ.
3.4 Results

Identification of two glycan binding sites in RII PfEBA-140

To identify receptor binding pockets, we solved the crystal structure of RII PfEBA-140 in complex with α-2,3-sialyllactose (Data collection and refinement statistics presented in Table 1). The glycan, α-2,3-sialyllactose, is a trisaccharide composed of sialic acid, galactose, and glucose and thus contains the critical sugar groups present on GPC that are required for binding (Neu5Ac(α-2,3)-Gal). RII PfEBA-140 bound to siallylactose is present as a monomer in the asymmetric unit, consistent with the solution and crystal structures of unbound RII PfEBA-140 (Lin et al., 2012). Clear density was observed for the sialic acid of two siallylactose molecules in the co-crystal structure, with one molecule contacting each DBL domain. The sialic acid moieties were modeled into the two binding sites (Fig. 1, A-D). Average B-factors for the sialic acids in both pockets were comparable (64.42 and 64.04 Å²), suggesting similar occupancy and affinity. Additional density was observed for the galactose of each sialyllactose molecule, but due to inherent flexibility galactose could not be adequately modeled. The galactose density extends out of both pockets and away from the contact surface (Fig. 1, B and C), supporting the importance of contacts between sialic acid and PfEBA-140. The two binding sites are present within a region spanning both DBL domains that contains a concentration of positively charged residues previously identified as a putative receptor binding region (Fig. 1D) (Lin et al., 2012). The glycan binding sites are in a structurally similar position on the F1 and F2 domains, with the sialic acid inserting into a valley at the interface of a helix and an extended loop (Fig. 1, E and F).

The sialic acid moiety contains a pyranose ring with single carboxyl, glycerol, and acetamido functional groups. The pyranose ring and functional groups of the bound sialic acid insert into
the F1 binding pocket and several contact residues form a cup-like interaction with the glycan (Fig. 2A). In the base of the pocket Gln182 hydrogen bonds and forms van der Waal's interactions with the sialic acid carboxyl. Contacts with the sialic acid glycerol chain are mediated by Arg158, which forms a hydrogen bond and is involved in van der Waals contacts with hydroxyl groups. On the opposite face of the binding pocket, the sialic acid acetamido group is oriented towards a helix containing Asn251. This helix places Asn251 in proper orientation to hydrogen bond with the 4-hydroxyl of the pyranose ring. In addition to the aforementioned contact surfaces, Ile181 and Ile185 form the base of the pocket and interact with hydrophobic regions of the sugar.

The F2 binding site contains a glycan contact surface that is similar to that observed in the F1 domain (Fig. 1F). However, the pocket residues are altered such that the observed glycan contacts are distinct (Fig. 2B). Asn457 forms contacts not observed in the F1 pocket as a result of its location on a loop at the top of the binding site. This orientation allows Asn457 to form van der Waal's interactions with the glycan carboxyl. Interactions with the sialic acid glycerol chain are mediated by Lys464, which hydrogen bonds with a hydroxyl group and engages in van der Waal's contacts with the sugar. Thr487 is present in the base of the pocket and contacts hydrophobic regions of the glycan. A helix containing Tyr556, analogous to the helix carrying Asn251 in the F1 pocket, extends the side chain of Tyr556 towards the glycan to engage in hydrophobic contacts with the sialic acid acetamido group.

**Distinct functional roles for each glycan binding pocket**

To examine the receptor specificity of the RII PfEBA-140 construct used for crystallization, we assessed the binding phenotype with enzyme treated erythrocytes. The rosetting interaction is sensitive to neuraminidase and trypsin treatment, but resistant to chymotrypsin, consistent with
previously observed results for RII PfEBA-140 (Fig. 2C) (Mayer et al., 2001, Narum et al., 2002, Maier et al., 2009). These results confirm that the interaction is sialic acid dependent and support the specificity of the recombinant construct for GPC on the erythrocyte surface.

To determine the role of specific sialic acid interactions identified in the crystal structure, individual glycan contact residues were mutated to alanine and the mutant constructs were tested for erythrocyte binding. Mutating individual contact residues in the F1 pocket severely diminished erythrocyte binding, suggesting that F1 glycan contacts are critical for receptor recognition (Fig. 2, D and F). In the absence of contacts mediated by Arg158 or Gln182, erythrocyte binding does not occur, suggesting that these residues are essential for binding interactions. The helix that presents the side chain of Asn251 is supportive for binding and it is likely that this helix plays an important role in maintaining the structure of the binding pocket. The hydrophobic base of the pocket, composed primarily of Ile181, is crucial for erythrocyte binding. The observed binding deficiency upon mutating Ile181 to alanine is likely due to its role in maintaining normal binding site conformation and engaging in strong hydrophobic contacts with the glycan.

In contrast to the severe phenotypes observed for alanine mutations in the F1 binding pocket, mutations in the F2 binding site resulted in a modest reduction in erythrocyte binding (Fig. 2, E and G). Mutation of Lys464 to alanine resulted in a 9.3% decrease in binding while mutation of other residues to alanine in the F2 pocket did not alter the erythrocyte binding phenotype. However, it is possible that a single mutation to alanine is insufficient to disrupt the function of this binding site. To further examine the functional role of glycan contacts in the F2 binding pocket, two residues were mutated to glutamate to induce charge repulsion with the sialic acid. Individually mutating Lys464 or Tyr556 to glutamate resulted in a 30.6% and 32.5% decrease in
binding, respectively. The F2 binding pocket is also close to residues in a basic patch (Arg485 and Asp554) that when mutated individually to alanine drastically reduce binding to erythrocytes (see Fig. 2B) (Lin et al., 2012). Together, these results suggest that the F2 binding site recognizes sialic acid in a functional, but non-essential manner and that the F2 sialic acid binding pocket and surrounding residues contact GPC. Reduced erythrocyte binding of alanine and glutamate mutants was not due to inefficient surface expression of RII PfEBA-140 as expression was monitored by GFP (Fig. 2, C-E). In addition, Lys347 was mutated to alanine as a surface control and displays wild type binding.

**PfEBA-140 glycan binding contacts are unique among sialic acid binding proteins**

The glycan contacts observed in the two PfEBA-140 binding pockets are unique when compared to other proteins that require cell surface sialic acids for binding. In addition to *Plasmodium* species, other Apicomplexan parasites, including *Toxoplasma gondii*, actively invade host cells during infection. Invasion by *T. gondii* is mediated by micronemal proteins (MICs). A member of this group, TgMIC1, recognizes sialic acid derivatives on the cell surface (Blumenschein et al., 2007). Glycan recognition by TgMIC1 requires contact with the sialic acid carboxyl group mediated by a threonine conserved in the TgMIC1 protein fold (Fig. 3A). Recombinant TgMIC1 does not bind host cells when this threonine is mutated to alanine, confirming the essential role of this residue (Blumenschein et al., 2007). In addition to the threonine, a histidine interacts with the carboxyl group and a tyrosine forms the base of the sialic acid binding pocket. The overall structure of the sialic acid binding site and the critical contact residues are thus distinct from both binding sites of PfEBA-140 (Fig. 3A).

In addition to Apicomplexan parasites, other pathogens recognize cell surface sialic acids during infection. All subtypes of the Influenza virus bind to host cells in a sialic acid dependent
manner mediated by the surface protein hemagglutinin. The conserved hemagglutinin sialic acid binding site utilizes essential glycerol chain contacts and hydrophobic interactions with the acetamido group to engage the glycan (Fig. 3B). Contacts with the carboxyl appear to be supportive (Watowich et al., 1994, Skehel et al., 2000, Ha et al., 2001). In comparison, although glycerol contacts are essential to PfEBA-140 binding, the binding site structures and interacting residues are distinct. In addition, interactions with the sialic acid carboxyl are critical for PfEBA-140, but not hemagglutinin binding.

Sialic acid is also required for a number of interactions involving proteins that modulate immune function, including selectins and Siglecs (Varki et al., 2007). Selectins interact most extensively with the sialic acid carboxylate and glycosidic oxygen and also contact the 4-hydroxyl of the pyranose ring (Fig. 3C). Contact with other functional groups does not appear to be important for selectin binding (Somers et al., 2000). This mode of contact sharply contrasts with that of PfEBA-140, which interacts with each functional group and requires both the carboxyl and glycerol moieties to effectively engage the glycan. The PfEBA-140 binding sites are also distinct from that of the Siglecs. Siglec binding is dependent on a salt bridge formed with the sialic acid carboxyl group and van der Waal’s contacts with the acetamido methyl group, mediated by a conserved arginine and semiconserved tryptophan/tyrosine, respectively (Fig. 3D) (Attrill et al., 2006, Crocker et al., 2007). Although other sialic acid binding proteins utilize contacts that resemble those observed in RII PfEBA-140, the GPC binding sites are distinct and represent a novel mode of glycan contact.

**Structural basis for altered receptor-binding caused by polymorphic residues**

Polymorphisms in PfEBA-140 are thought to reduce affinity to GPC (Maier et al., 2009) and have been proposed to alter PfEBA-140 receptor specificity (Mayer et al., 2002). One of the
identified polymorphisms, Ile185, is buried in the critical F1 binding pocket (see Fig. 2A). Variants that contain a valine in place of isoleucine at position 185 display reduced binding to erythrocytes and GPC (Maier et al., 2009). The F1 binding pocket forms a tight fit with high shape complementarity to the sialic acid when isoleucine is present at position 185 (Fig. 4A). To examine the effect of a valine mutation, Ile185 was mutated to valine in silico. This mutation opens a cavity at the base of the pocket adjacent to the 4-hydroxyl of the pyranose ring (Fig. 4B). Mutation to valine in vivo likely alters the pocket surface in this manner, reducing shape complementarity and increasing solvent accessibility to both the pocket and the glycan. This change would disrupt glycan binding and provides a structural basis for diminished affinity of polymorphic PfEBA-140 variants for erythrocytes (Maier et al., 2009). The structural alteration also elucidates a putative basis for altered receptor specificity.

**Glycan binding induces a helix shift in the F2 domain**

The structure of RII PfEBA-140 bound to sialyllactose maintains the same overall architecture and hinge angle as unbound RII (Fig. 5A). Glycan binding did not result in any large structural changes in the F1 domain, but caused a helix containing Thr493, Tyr494 and Leu495 to move away from the binding pocket of the F2 domain (Fig. 5B). The bound sialic acid in the F2 domain inserts into a cavity containing two water molecules in the unbound structure. Glycan binding blocks access of the water molecules at this position, disrupting main chain contacts and destabilizing the helix such that the helix movement is possible (Fig. 5C). The glycan induced movement in the F2 domain may provide access to the highest affinity interface for GPC.

**Putative sulfate-binding motifs map close to the glycan binding pockets**

In addition to binding GPC, RII PfEBA-140 is capable of interactions with heparin and potentially with cell surface GAGs. Heparitinase treatment of erythrocytes inhibits RII PfEBA-
140 binding, providing support for a potential interaction with cell surface GAGs such as heparan sulfate (Kobayashi et al., 2010). These observations suggest that GAGs on the erythrocyte surface interact with PfEBA-140 and may promote invasion. RII PfEBA-140 contains five putative sulfate-binding motifs that can potentially bind GAGs, three of which are adjacent to the identified sialic acid binding sites in F1 and F2 (Fig. 5D). Mutation of the adjacent sulfate-binding motifs 1 and 4 greatly diminishes erythrocyte binding. The effect of mutating motif 2 has not been examined. In contrast to the adjacent motifs 1 and 4, mutation of sulfate-binding motifs 3 and 5, which are distal to the glycan binding sites, has little or no effect on erythrocyte binding (Kobayashi et al., 2010). Thus, it is plausible that sulfate-binding sites 1, 2, and 4 represent true interaction elements that may play a role in facilitating invasion. While the putative sulfate-binding motifs are proposed to engage heparan sulfate (Kobayashi et al., 2010), additional sulfate containing molecules such as sulfatides, sulfated glycolipids, glycoprotein oligosaccharides, and chondroitin sulfate proteoglycans exist on erythrocytes. It is possible that putative sulfate-binding motifs 1, 2, and 4 may interact with any or all of these moieties.


3.5 Discussion

The results presented here represent the first molecular description of specific interactions an EBL protein utilizes to recognize and bind sialic acid. Although DBL domains have been crystallized with sugars previously, the crystal structure of RII PfEBA-140 presented here represents the first full modeling of sialic acid into a DBL domain sugar binding site. This modeling allowed identification of two receptor binding sites and elucidation of critical molecular interactions that the parasite utilizes to form the essential tight junction with the host erythrocyte membrane during invasion. Mapping molecular contacts demonstrated that PfEBA-140 recognizes sialic acid on receptor glycans utilizing unique interactions compared to its paralog PfEBA-175 as well as compared to viral cell invasion and host immune proteins that also engage sialic acid. The bound structure also provided insight into the role of polymorphisms in altering receptor binding and illuminated the location of putative sulfate-binding motifs that may engage GAGs or other sulfate-containing surface molecules. The importance of these moieties during invasion is unclear, but localizing their binding sites within RII provides a basis for future studies examining their function during Plasmodium invasion.

The identification and description of receptor binding sites within PfEBA-140 allows for direct comparison with the glycan binding sites of other DBL domains. DBL domains function in a number of settings including erythrocyte invasion and cytoadherence of infected erythrocytes mediated by Erythrocyte Membrane Protein 1 (PfEMP1) (Kraemer et al., 2006). The two glycan binding regions of RII PfEBA-140 are structurally distinct from receptor-binding regions of PfEBA-175, PvDBP, PfEMP1 DBL3X and the PfEMP1 VarO head region, and thus represent novel receptor-binding sites for the DBL fold (Tolia et al., 2005, Singh et al., 2008, Batchelor et al., 2011, Vigan-Womas et al., 2012). Specifically, the glycan interactions observed in RII
PfEBA-175 are limited to the dimer interface (Tolia et al., 2005). In contrast, formation of the glycan binding sites in RII PfEBA-140 does not require dimerization (Fig. 6). We previously demonstrated that RII PfEBA-140 is monomeric in solution in the absence of receptor (Lin et al., 2012). RII PfEBA-140 has been observed only as a monomer in solution and crystal forms, and the lack of dimerization in the presence of sialyllactose suggests that PfEBA-140 may engage GPC as a monomer. Studies examining the oligomeric state of PfEBA-140 in the presence of GPC are required to fully understand the mechanism of receptor engagement and specificity. In addition to the different oligomeric states observed upon sialic acid binding for PfEBA-175 and PfEBA-140, the location within the DBL domain and overall structure of the binding pockets is distinct (Fig. 6). Furthermore, PfEBA-175 contains six glycan binding sites per dimer, in contrast to two observed for the monomer of PfEBA-140. It is likely that the novel modes of contact observed for PfEBA-140 confer specificity for sialic acid moieties present on GPC over Glycophorin A or B, and suggests how structurally similar DBL domains specifically recognize diverse receptors.

In addition to structural and molecular contrasts with the glycan binding sites of DBL domains, this work allows for comparison with other Apicomplexan invasion proteins that recognize cell surface sialic acid. As shown in Figure 3A, sialic acid recognition by TgMIC1 is quite distinct when compared to PfEBA-140. In addition, there is no structural similarity in the overall protein fold of the minimal binding regions of TgMIC1 and PfEBA-140 (Blumenschein et al., 2007). However, both proteins contain two tandem, conserved domains comprising the minimal binding region, with each domain containing a single sialic acid binding site on the same face of the protein. This resemblance suggests some functional similarities in their mode of glycan engagement (Blumenschein et al., 2007). The Plasmodium merozoite is specifically
marked for the erythrocyte, unlike *T. gondii*, which is capable of invading any nucleated cell. It thus appears that TgMIC1 has developed a sialic acid binding mode that allows recognition a broad range of glycans during cell binding, sharply contrasting the proposed erythrocyte sialic acid glycan specificity of PfEBA-140. PfEBA-140 receptor specificity may thus be conferred by a unique glycan composition on the erythrocyte surface combined with essential interactions with the GPC protein backbone.

The erythrocyte binding phenotypes observed upon mutation of contact residues in each glycan pocket suggest distinct roles for the individual DBL domains of RII PfEBA-140 during erythrocyte binding. GPC contains several O-linked glycans and a single N-linked glycan, each with the essential sugar moiety Neu5Ac(α-2,3)-Gal. Recognition of the N-linked glycan is essential to the interaction (Mayer *et al.*, 2006). Thus, interactions with the N-linked glycan may be mediated solely by the F1 sialic acid binding site. The null and severely deficient binding phenotypes observed upon mutation of contact residues in the F1 pocket could be the result of losing the critical interaction with the N-linked glycan.

Receptor contacts with the F2 domain are also required for erythrocyte binding as neither individual DBL domain is sufficient to bind erythrocytes independently (Mayer *et al.*, 2002, Lin *et al.*, 2012). The observation that only mutations to glutamate in the F2 binding pocket reduced binding efficiency may be explained by compensatory interactions between the glycan and other binding site residues. In addition to glycan contacts, PfEBA-140 erythrocyte binding is dependent on the GPC protein backbone. It is thus also plausible that F2 engages in strong contacts with additional regions of GPC and that glycan interactions are supportive but not essential in maintaining this interaction.

In addition to invading via GPC, it has been proposed that polymorphic variants of PfEBA-
140 are capable of invading using an alternative receptor. The identification of Ile185 in the F1 binding pocket suggests that polymorphism at this residue may play an important role in determining receptor affinity and possibly specificity. Our studies suggest a model to explain the altered binding profile of polymorphic variants. The proximity of the glycan binding pockets and sulfate-binding motifs 1, 2 and 4 suggest that GPC binding precludes GAGs or sulfate-containing molecules from accessing these sugar binding sites. This proposal is supported by the fact that the helix containing sulfate-binding motif 4 is shifted upon sialyllactose binding (see Fig. 5, A and B). It is plausible that a valine mutation at position 185 destabilizes the F1 binding pocket, reducing affinity for sialic acid and opening sulfate-binding motifs 1 and 2 for functional GAG binding. The altered surface observed following in silico mutation of Ile185 to valine supports the destabilization hypothesis (see Fig. 4). It is equally possible that a secondary sialylated receptor is recognized by specific variants of PfEBA-140 during invasion.

The structural and functional data described here have provided a detailed molecular description of sialic acid recognition by an EBL family erythrocyte invasion ligand, PfEBA-140. Preventing erythrocyte receptor engagement provides an excellent opportunity to inhibit merozoite invasion and blood stage growth of the parasite. The identification of critical interactions at the invasion interface will thus aid in the design of rational therapeutics and vaccines that target the EBL ligands.
3.6 Figures and Tables

Figure 1. The crystal structure of RII PfEBA-140 bound to sialyllactose reveals two receptor glycan binding sites. A, RII PfEBA-140 bound to two sialyllactose molecules in the crystal structure. The F1 and F2 DBL domains each form a contact surface with one sialyllactose molecule. The F1 domain is shown in orange, the F2 domain in blue. A short linker connecting the two DBL domains is shown in grey. The modeled sialic acid molecules in each glycan contact site are shown in green and boxed in black. B, The bound sialic acid inserts into the F1 and F2 binding pockets while the galactose extends away from the pocket surface. The sialic acid molecules and corresponding electron density observed in the F1 domain (left, orange) and F2 domain (right, blue) are displayed within their binding sites. The sialic acid is shown in red for clarity. Density for the galactose is shown to illustrate the location of this sugar moiety relative to the sialic acid binding site. The Fo-Fc map prior to sialic acid modeling is shown in green and contoured at 3.0σ. The 2Fo-Fc map obtained following modeling of the sialic acid into the binding pocket is shown in blue and contoured at 1.0σ. C, Close-up view of the electron density observed for the sialic acid and galactose of the bound sialyllactose in the F1 (left) and F2 (right) domains. The density maps are contoured and colored as in B. D, The two glycan binding sites are located in a region of concentrated positively charged residues that may function as a high affinity interface for GPC engagement. Surface charge potential is colored from +3.5eV (blue) to -3.5eV (red). E, The sialic acid binding pockets are in structurally similar locations on the F1 and F2 DBL domains. The glycan bound to the F1 domain is shown in green, the glycan bound to the F2 domain is shown in yellow. F, The individual binding sites in each DBL domain are structurally similar. Specifically, the sialic acid inserts into a valley at the interface of a helix and an extended loop present on both domains. The glycan coloring is the same as in E.
Figure 2. Sialic acid modeling and erythrocyte binding studies provide a detailed molecular description of the RII PfEBA-140 erythrocyte invasion interface. A, Close-up view of the F1 binding pocket. Four residues directly contact the sialic acid. Hydrogen bonds are designated with black lines. B, Close-up view of the F2 binding pocket. Four residues directly contact the sialic acid. The single hydrogen bond interaction is shown with a black line. C, Erythrocyte binding by RII PfEBA-140 expressed on the surface of HEK-293 cells is sensitive to trypsin and neuraminidase treatment of red blood cells, but chymotrypsin resistant. Erythrocytes were treated with 0.1 mg/mL trypsin, 0.1 mg/mL chymotrypsin, or 5 mU of *Vibrio cholera* neuraminidase at 37°C prior to the rosetting assay. Bound erythrocytes appear black around the transfected mammalian cells. The images are displayed at 20X magnification. D and E, RII PfEBA-140 mutants were expressed on the surface of mammalian cells and tested for erythrocyte binding. GFP was used to assess proper expression of each construct. Wild type RII PfEBA-140 extensively binds erythrocytes, and a construct containing GFP alone is not capable of binding. To identify critical binding interactions, individual glycan contact residues in the F1 and F2 binding pockets were mutated to alanine or glutamate and tested for erythrocyte binding by rosetting assay. Bound erythrocytes appear black around the transfected mammalian cells. The images are displayed at 20X magnification. F and G, Percentage of cells expressing point mutants of RII PfEBA-140 that bind erythrocytes relative to wild type RII. The binding phenotypes are representative of the binding percentage quantified for 30 fields of view per mutation. Lys347, mutated to alanine as a surface control residue, displays wild type binding.
Figure 3. RII PfEBA-140 glycan binding sites are distinct from other classes of sialic acid binding proteins. The PfEBA-140 F1 (orange) and F2 domains (blue) are shown overlayed with the A, TgMIC1 (light blue), B, Hemagglutinin (red), C, Selectin (purple), and D, Siglec (light green) sialic acid binding sites. The sialic acid molecule bound to PfEBA-140 is shown in green and the sialic acid bound to the protein of comparison is shown in grey. Arrows highlight structural differences in each case.
Figure 4. The F1 binding pocket surface is altered when Ile185 is mutated to valine in silico. A, The glycan binding surface containing Ile185 (the residue present in the construct used for crystallization) exhibits strong shape complimentary with the sialic acid. The overall surface is shown in grey on the top, the electrostatic surface is shown on the bottom (surface charge potential is colored from +3.5eV (blue) to -3.5eV (red)). The bound sialic acid is shown in green. B, The binding pocket surface is altered following in silico mutation of residue 185 to valine. The top displays the overall surface of the pocket in grey, and the bottom displays the electrostatic surface (surface charge potential is colored as in A). The altered surface cavity is identified with an arrow and the sialic acid is shown in green.
FIGURE 5. The sialyllactose bound structure illuminates a glycan induced structural change and allows for localization of putative sulfate-binding sites. A, Overlay of the sialyllactose bound (F1 domain in orange, F2 domain in blue) and unbound (F1 domain in light orange, F2 domain in light blue) structures of RII PfEBA-140. The bound sialic acid molecules are shown in green. The glycan induced structural change is outlined in black. B, Close-up view of residues Thr493, Tyr494, and Leu495, which show a pronounced movement following glycan binding. C, The helix shift in the F2 domain observed upon glycan binding is propagated by sialic acid interference with main chain water molecule interactions. Glycan binding precludes access of two water molecules to the cavity indicated by the black arrow (the waters are present in the apo structure). In the absence of contacts mediated by these water molecules, the helix is destabilized, allowing the movement observed in the bound crystal structure. The water molecules are shown in cyan and the sialic acid in green. The bound F2 domain is shown in dark blue; the unbound F2 domain in light blue. D, Three sulfate-binding motifs in RII PfEBA-140 are adjacent to the sialic acid binding sites. Two other putative sulfate-binding motifs are distal to the binding sites. The putative sulfate-binding sites are shown in purple and identified with an arrow.
Figure 6. PfEBA-140 sialic acid recognition is distinct from that of PfEBA-175. Shown here is an overlay of the RII PfEBA-140 monomer with the dimer of RII PfEBA-175. The PfEBA-140 F1 domain is shown in orange and the F2 domain is shown in blue. The RII PfEBA-175 monomer overlayed with RII PfEBA-140 is shown in light green. The second monomer of RII PfEBA-175 is shown in light purple. Residues of PfEBA-140 involved in sialic acid binding are shown in red. The sialic acid molecules bound to RII PfEBA-140 are shown in black for clarity.
Table 1: Data collection and refinement statistics

<table>
<thead>
<tr>
<th></th>
<th>PfEBA-140 RII sialyllactose complex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data collection</td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
</tr>
<tr>
<td>(a, b, c) (Å)</td>
<td>64.95, 76.03, 81.43</td>
</tr>
<tr>
<td>(\alpha, \beta, \gamma) (°)</td>
<td>90.00, 96.93, 90.00</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>20-3.0 (3.09-3.0)</td>
</tr>
<tr>
<td>(R_{\text{merge}})</td>
<td>7.8% (65.5%)</td>
</tr>
<tr>
<td>(I/\sigma I)</td>
<td>14.06 (2.05)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>96.2% (97.7%)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.00 (3.00)</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>20.0-3.0</td>
</tr>
<tr>
<td>No. reflections</td>
<td>15,336</td>
</tr>
<tr>
<td>(R_{\text{work}} / R_{\text{free}}) (%)</td>
<td>21.82 / 25.48</td>
</tr>
<tr>
<td>No. atoms</td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>5,266</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>42</td>
</tr>
<tr>
<td>Water</td>
<td>22</td>
</tr>
<tr>
<td>(B)-factors</td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>87.18</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>64.23</td>
</tr>
<tr>
<td>Water</td>
<td>42.74</td>
</tr>
<tr>
<td>r.m.s.d</td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.001</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>0.437</td>
</tr>
<tr>
<td>Ramachandran</td>
<td></td>
</tr>
<tr>
<td>Favored</td>
<td>95.52%</td>
</tr>
<tr>
<td>Allowed</td>
<td>4.48%</td>
</tr>
<tr>
<td>Disallowed</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

Values in parentheses are for highest-resolution shell.
3.7 References


Chapter 4

PvDBP utilizes unique receptor binding interfaces and a novel two-step binding process to engage its RBC receptor
Preface

The work presented here was performed by me, Joseph D. Batchelor, Natalie S. Omattage, Gregory T. Dekoster, Katherine Henzler-Wildman, and Niraj H. Tolia. Crystallization, data collection and structure solution was performed by J.D.B. NMR experiments were performed by J.D.B. and G.T.D. I performed the ITC experiments and analyzed the resulting data that supported the two-step binding mechanism initially observed in the crystal structures of PvDBP in complex with DARC. I also analyzed the structures and performed functional RBC binding studies that illuminated key amino acids that determine receptor specificity among PvDBP homologues in *P. knowlesi*. The manuscript was written predominantly by J.D.B. and N.H.T. I wrote portions of the manuscript and prepared figures related to ITC experiments, RBC binding experiments, and put together tables that describe the complete amino acid composition of receptor binding sites. I also completed the peer review process and responded to all critiques and comments related to experiments and the written manuscript. This work was published in its entirety: [Batchelor J.D., Malpede B.M., Omattage N.S., DeKoster G.T., Henzler-Wildman K.A., and Tolia, N.H. (2014) Red Blood Cell Invasion by *Plasmodium vivax*: Structural Basis for DBP Engagement of DARC. *PLoS Pathogens* 10(1): e1003869].
Acknowledgements

We thank S. Beverley and D. Sibley for constructive comments on the manuscript, and T. Lohman, A. Kozlov, D. Fremont, G. Amarasinghe, J. Binning, and T. Brett for assistance with isothermal titration calorimetry. We thank J. Nix and ALS Beamline 4.2.2 (contract DE-AC02-05CH11231); S. Ginelle and the APS SBC-CAT beamline 19-ID (contract DE-AC02-06CH11357).
4.1 Abstract

*Plasmodium* parasites use specialized ligands which bind to red blood cell (RBC) receptors during invasion. Defining the mechanism of receptor recognition is essential for the design of interventions against malaria. Here, we present the structural basis for Duffy antigen (DARC) engagement by *P. vivax* Duffy binding protein (DBP). We used NMR to map the core region of the DARC ectodomain contacted by the receptor binding domain of DBP (DBP-RII) and solved two distinct crystal structures of DBP-RII bound to this core region of DARC. Isothermal titration calorimetry studies show these structures are part of a multi-step binding pathway, and individual point mutations of DARC contact residues result in a complete loss of RBC binding by DBP-RII. Two DBP-RII molecules sandwich either one or two DARC ectodomains, creating distinct heterotrimeric and heterotetrameric architectures. The DARC N-terminus forms an amphipathic helix upon DBP-RII binding. These structures of a *Plasmodium* ligand in complex with a red blood cell receptor define the molecular details of engagement for a pervasive class of ligands containing the Duffy binding-like fold. The studies reveal a receptor binding pocket in DBP and critical contacts in DARC, reveal novel targets for intervention, and suggest that targeting the critical DARC binding sites will lead to potent disruption of RBC engagement as complex assembly is dependent on DARC binding. These results allow for models to examine inter-species infection barriers, *Plasmodium* immune evasion mechanisms, *P. knowlesi* receptor-ligand specificity, and mechanisms of naturally acquired *P. vivax* immunity. A step-wise binding model is proposed that may facilitate signaling through the transmembrane and cytoplasmic domains of DBP to activate pathways of invasion. It is anticipated that the structural basis of DBP host-cell engagement will enable development of rational therapeutics targeting this interaction.
4.2 Author Summary

Malaria parasites, including *Plasmodium vivax*, must actively invade erythrocytes during blood stage growth in humans. *P. vivax* Duffy Binding Protein (DBP) is a critical invasion ligand that recognizes the receptor Duffy antigen/Receptor for chemokines (DARC) during invasion. To identify critical binding contacts during parasite red blood cell invasion and determine the molecular basis of DBP receptor recognition, we identified the minimal region of DARC contacted by DBP and performed structural studies on the minimal binding domain of DBP in complex with the minimal region from DARC. These studies revealed that two DBP molecules bind two DARC molecules. We performed erythrocyte binding assays with binding site mutants and identified essential receptor contacts. The identification of receptor binding sites and molecular interactions critical to the invasion process provides a basis for targeted disruption of erythrocyte invasion mediated by DBP. The structural and functional studies of DBP and DARC presented here may aid in the rational design of vaccines and invasion inhibitory therapeutics.
4.3 Introduction

*Plasmodium vivax* is a widely distributed human parasite, with 40% of the world’s population at risk of infection and an estimated 70-130 million cases of *P. vivax* malaria each year (Mendis *et al.*, 2001, Price *et al.*, 2007). *P. vivax* is prevalent in India, Southeast Asia, and South America (Mendis *et al.*, 2001), but is rare in most of Sub-Saharan Africa (Welch *et al.*, 1977). This rarity is the result of a silencing mutation in the Duffy blood group, found at frequencies near fixation in Sub-Saharan Africa (Howes *et al.*, 2011), that confers resistance to *P. vivax* (Miller *et al.*, 1976). This phenotype has arisen independently at least three times, and *P. vivax* in malaria endemic regions has driven selection for the Duffy negative phenotype. This phenotype confers protection against *P. vivax* because during red blood cell (RBC) invasion the *P. vivax* Duffy Binding Protein (DBP) binds the Duffy antigen/receptor for chemokines (DARC) on RBCs (Wertheimer *et al.*, 1989, Horuk *et al.*, 1993). Therefore, RBCs which lack DARC are refractory to *P. vivax* invasion. DARC is an atypical GPCR, thought to serve as a ‘reservoir’ for excess inflammatory chemokines (Gardner *et al.*, 2004).

Repeated cycles of RBC invasion and rupture cause the clinical symptoms of malaria. To invade a RBC, *Plasmodium* merozoites release the contents of specialized apical organelles: the micronemes and rhoptries. DBP is a member of the erythrocyte binding-like (EBL) family of proteins, which localize to micronemes and use Duffy binding-like (DBL) domains to bind specific RBC receptors with high affinity. DBL domains are located in “region II” of EBL proteins (Chitnis *et al.*, 1994, Tolia *et al.*, 2005), and DBP region II (DBP-RII) is required for formation of a tight junction between *Plasmodium* and RBC membranes. DBP is an exceptional *P. vivax* therapeutic target because it is the sole EBL family member in the *P. vivax* genome.
This contrasts with *P. falciparum*, which has multiple, redundant EBL family members which mediate RBC invasion by binding different host RBC receptors.

DBP is a leading vaccine candidate against *P. vivax* malaria (Beeson *et al.*, 2007). Individuals living in endemic regions develop natural immunity to *P. vivax* in an age-dependent manner which strongly correlates with humoral and cellular recognition of DBP-RII (Xainli *et al.*, 2002, Xainli *et al.*, 2003, Cole-Tobian *et al.*, 2009). Antibodies against DBP inhibit *P. vivax* RBC invasion (Grimberg *et al.*, 2007), and antibody epitopes in DBP-RII recognized by inhibitory antibodies have been identified (Chootong *et al.*, 2010). However, due to a high level of polymorphism in DBP-RII and the selection for strain-specific immunity, identifying residues that are essential to the invasion interaction is still a critical step towards defining vaccination targets.

Previous studies have illuminated key determinants of DBP-RII binding to DARC and begun to define essential elements of the binding interaction. DARC exists as two codominant alleles, Fya and Fyb, with a single polymorphism at residue 42. Fya contains a glycine, and Fyb an aspartate (G/D42) at this position. The Fyb phenotype has been shown to increase binding to DBP-RII (King *et al.*, 2011). It is also known that DARC is sulfated on the erythrocyte surface at tyrosine residues 30 and 41, and sulfation of tyrosine 41 has been shown to play a role in binding to DBP (Choe *et al.*, 2005). Specifically, a sulfated recombinant DARC N-terminus construct inhibits the DBP-RII erythrocyte interaction to a greater extent than an unsulfated construct. Extensive functional studies have also suggested interaction residues for both DBP-RII and DARC (Chitnis *et al.*, 1996, Singh *et al.*, 2003, VanBuskirk *et al.*, 2004, Hans *et al.*, 2005, Mayor *et al.*, 2005, Tournamille *et al.*, 2005). The crystal structure of DBP-RII has been solved (Batchelor *et al.*, 2011), and illuminated a putative sulfotyrosine binding pocket. Biophysical
studies have demonstrated that a non-sulfated DARC construct functionally binds and is capable of inducing dimerization of DBP-RII (Batchelor et al., 2011), suggesting that regions outside of the sulfotyrosine residues play an important role in the binding interaction. Despite these studies on the DBP-RII interaction with DARC, the full mechanism of binding and complete extent of molecular interactions between these binding partners are not fully understood.

In an effort to define the mechanism of DBP red blood cell binding and to identify specific molecular interactions at the *P. vivax* invasion interface, DBP-RII was crystallized with the DARC ectodomain. Two crystal forms were observed, a heterotrimeric complex in which one DARC molecule binds the DBP-RII dimer, and a heterotetrameric complex containing two DARC molecules bound to the DBP-RII dimer. The crystal structures of these two complexes represent the first structural characterization of a receptor-bound *Plasmodium* EBL ligand and provide insight into the structure of a portion of the DARC ectodomain that mediates this interaction. The structures illuminate DARC contact residues that explain inter-species barriers to *P. vivax* infection. In addition, point mutations in DARC binding site residues within *P. knowlesi* DBP homologs provide a potential model for the molecular basis of receptor specificity within the EBL family. The heterotetrameric complex shows that the DARC molecules are bound in parallel, resulting in a *Plasmodium* proximal face and a RBC proximal face of DBP-RII. The structures confirm the previously identified DBP-RII dimer interface (Batchelor et al., 2011) as a putative target of protective immunity, and reveal novel targets for naturally acquired immunity, including the RBC proximal face of DBP-RII and the DARC-binding pockets. Our studies also provide the basis for a framework defining the mechanism of DARC engagement by *P. vivax*. Specifically, the two bound forms of DBP-RII observed in the crystal structures suggest that an initial binding event followed by receptor-induced DBP dimerization leads to a
DBP:DARC heterotrimer that subsequently binds to a second DARC monomer to create the final heterotetrameric assembly. Isothermal titration calorimetry (ITC) experiments performed with recombinant DARC and DBP-RII support the two state, induced dimerization model of erythrocyte engagement.
4.4 Results

**DBP-RII engages the central region of the DARC N-terminal ectodomain**

DARC binds DBP-RII through its N-terminal 60 amino acid ectodomain (Horuk et al., 1993) and DARC binding induces dimerization of DBP-RII (Batchelor et al., 2011). Binding and dimer-induction *in vitro* has been demonstrated for non-tyrosine sulfated DARC N-terminal ectodomain (Batchelor et al., 2011). NMR experiments were performed to determine the region of an unsulfated version of the DARC N-terminal ectodomain contacted by DBP-RII (Fig. 1). Resonance assignments for the 60 N-terminal amino acids of DARC (DARC 1-60) were obtained by standard triple resonance experiments. The peaks are not well dispersed (Fig. 1) and chemical shifts closely match canonical random coil chemical shifts, consistent with a lack of secondary structure.

In the presence of DBP-RII, the large size of the 88 kDa DBP-RII:DARC 1-60 complex led to significant broadening of many peaks, preventing full assignment. However, comparison of the spectra between bound and unbound states revealed that signals corresponding to the first N-terminal 15-16 amino acids overlay well and have only modest line broadening. This result suggests that this region remained unstructured upon binding and does not directly contact DBP-RII. At the C-terminus, residues from 44-60 exhibit some line broadening or chemical shift changes, but this region, similar to that of the N-terminus, was still relatively unperturbed upon binding to DBP-RII. In contrast, peaks corresponding to the central region of DARC1-60 became significantly broadened, shifted, or disappeared in the bound complex. These results indicate that residues within the central region of the DARC ectodomain are highly perturbed upon interaction with DBP-RII and are thus most likely to directly contact DBP and form the minimal binding
domain. This result is consistent with a region from DARC sufficient for blocking RBC binding by DBP-RII (Chitnis et al., 1996).

**The DARC ectodomain forms a helix that binds the dimer interface of DBP-RII**

Screening for crystallization conditions of DBP-RII in complex with DARC ectodomain constructs designed around the central binding region resulted in two crystal structures of the DBP-RII:DARC complex. The first was a 1.95Å crystal structure of a 2:1 complex of DBP-RII:DARC16-43 (Table 1). In this structure, two DBP-RIIs (DBP1 and DBP2) bind a single DARC (DARC A) creating the heterotrimer (Fig. 2A) with a total buried surface area of 2241.8 Å². The second structure was a 2:2 complex of DBP-RII:DARC14-43 that was refined to 2.6Å (Table 1). In the second structure, two DBP-RIIs each bind two DARCs (DARC A and DARC B) creating two DARC binding sites. This architecture creates a heterotetramer (Fig. 2B), with a total buried surface area of 3628.6 Å². We postulated that the two structures represent snapshots in the assembly of the DBP-RII:DARC complex and may define structural changes during step-wise binding. Additionally, in the heterotrimer, the second DARC binding site is preformed to accept another DARC. In both crystal forms the two DBP-RII molecules are not identical and no higher order symmetry exists in the asymmetric unit. Therefore, DARC binding results in two distinct DBP-RII molecules in each asymmetric unit.

**DBP-RII interacts with DARC in a step-wise binding process**

We utilized ITC to examine the mechanism of DBP-RII:DARC engagement and assembly in solution. ITC is an excellent technique to unambiguously determine interaction stoichiometries and can be applied to examine step-wise and multi-state binding systems in solution. Titration of DARC1-60 into DBP-RII demonstrated that DARC binding occurs in a step wise assembly consistent with the crystallographic studies. A biphasic binding isotherm indicative of a two-state
assembly was observed (Fig. 3). The first binding event has a molar ratio of 0.5, expressed in monomers of DBP-RII, indicative of a 2:1 heterotrimeric complex of (DBP-RII)$_2$:(DARC1-60)$_1$.

The second binding event occurs at a molar ratio of 1 indicative of a 2:2 heterotetrameric complex or (DBP-RII)$_2$:(DARC1-60)$_2$. The data were fit to a two independent site binding model which suggested affinities of $2151\pm352$\,nM for the first binding event and $56.8\pm50.4$\,nM for the second binding event, consistent with high affinity binding. However, it should be noted that the two independent site binding model is not a perfect description of DBP-RII:DARC binding as receptor-induced dimerization is not included in the model. Therefore, exact affinity determination will require further work necessary to define all thermodynamic parameters of binding. While the exact affinity will require a more detailed fitting model, the stoichiometries determined and thus the observation of step-wise assembly in solution are not affected by the fitting model selected and are reliable. In summary, the crystallographic and ITC solution data presented here demonstrate a multi-step sequential binding mechanism involving DARC-induced assembly of DBP-RII.

**Identification of molecular interactions between DARC and DBP-RII**

The NMR studies indicate DARC1-60 lacking three cysteines is unstructured in the absence of DBP-RII. In both structures, clear density is seen for DARC residues 19-30 (Fig. S1). DARC is induced to form an amphipathic helix in the crystal structures upon binding and engages a positively charged groove at the DBP-RII dimer interface (Fig. S1). All DARC interacting residues and the dimer interface of DBP-RII are located in subdomain 2 of DBP-RII. In addition to the dimer interface, each DARC binding site, one in the heterotrimer and two in the heterotetramer, can be broken into two interfaces: a primary DARC binding interface with one DBP-RII monomer, and a secondary DARC binding interface created by the second DBP-RII
monomer (Fig. 4 and 5, Table 2 and 3). The DBP-RII homodimer interface is composed of DBP1 residues I265-R274 and DBP2 residues F261-Y278 in the heterotrimer (Fig. 4C), and DBP1 residues H262-R274 and DBP2 residues F261-Y278 in the heterotetramer (Fig. 5B). The primary DARC binding interface in both structures consists of DBP-RII residues L270-K289 of helix 4 and Q356-K367 of helix 7 (Fig. 4 and 5). DBP-RII binds the amphipathic DARC helices through a hydrophobic core flanked by electrostatic interactions.

The secondary DARC binding interface is formed by residues V254 to F267 (loop 254-267) (Fig. 4B, 5C-D). When DBP-RII is not bound to DARC this region is disordered (Batchelor et al., 2011). Loop 254-267 contains residues which are required for binding (VanBuskirk et al., 2004), and are recognized by neutralizing antibodies (Chootong et al., 2010). When DARC is bound, loop 254-267 becomes ordered and engages the DARC bound by the primary interface of a neighboring DBP-RII. In the heterotrimer, residues H262-T266 make contacts at the secondary interface (Fig. 4B). In the heterotetramer, DBP1 residues R263-I265 and DBP2 residues F261-D264 contact DARC of the opposing monomer (Fig. 5C-D). Thus, DARC is sandwiched between two DBP-RII molecules in each DARC binding site.

**Architectural transitions upon receptor binding**

In the absence of receptor, DBP-RII crystallized as a dimer stabilized by phosphates (Batchelor et al., 2011). Although this prior structure resembles the receptor-bound conformation presented here, there are substantial structural differences in the architecture of the dimer compared to the heterotrimer or heterotetramer. These differences are crucial towards correctly defining the invasion interaction. In the heterotrimer structure, a new DBP-RII homodimer interface is created by a translation of 7 residues covering 12Å along helix 4, relative to the unbound structure (Fig. S2A). The heterotetramer structure has a larger translation along the
same interface in the same direction, with a second 12 Å displacement relative to the 
heterotrimer (Fig. S2B), and a 23 Å displacement relative to the unbound DBP-RII homodimer 
interface (Fig. S2C). The directionality of these transitions is consistent with sequential steps in a 
stepwise mechanism of receptor binding. While there are major changes in the DBP-RII:DARC 
complex architectures, individually each DBL domain aligns well to the DBP-RII DBLs solved in 
the absence of receptor (Fig. S2D-G).

**Critical contacts in DARC**

The observation that DARC 19-30 is contacted by DBP-RII is consistent with alanine 
scanning work (Tournamille *et al.*, 2005). Mutation of DARC residues 20-22 and 24-26 
abrogated binding in a direct protein interaction assay. Each of these residues, with the exception 
of D21, makes direct contacts with DBP-RII and are buried in the complex (Fig. 4 and 5). D21, 
which is required for DBP-RII binding but does not directly contact DBP-RII, stabilizes the 
DARC N-terminal helix dipole by positioning its acidic side chain directly over the helix. E23, 
on the other hand, is on the surface of the complex and is solvent exposed. Mutation of E23 to 
alanine had no effect on binding consistent with its location in the complex.

Sulfo tyrosine residues in DARC increase DARC’s ability to inhibit DBP-RII RBC binding 
(Choe *et al.*, 2005). A previous structure of DBP-RII alone identified a potential sulfo tyrosine 
binding site that includes residues K273 and Q356 (Batchelor *et al.*, 2011). In the receptor-bound 
structure presented here, the hydroxyl of DARC Y30 points directly at this pocket created by 
K273, and Q356 (Fig. S3). Therefore, sulfo tyrosine 30 appears to bind at this site when DARC is 
sulfated. DARC residues 14-43 were included in our crystallographic studies. However, clear 
density was only observed for residues Q19 to Y30, and no density was present for the remainder
of DARC. The crystallographic data along with the RBC binding studies discussed above support that residues 19-30 constitute a critical interaction site with DBP-RII.

**Residues in DBP-RII that contact DARC are required for RBC engagement**

Having identified residues of DBP-RII that directly contact DARC, structure-guided mutagenesis was used to determine whether this model of binding explains DBP engagement of RBCs (Fig. 6). Mutation of residues Y363 or A281 in the primary DARC binding interface led to a complete loss of binding in a functional RBC binding assay. This is expected as DARC binding drives complex formation and mutations preventing DARC binding will completely abrogate complex formation and attachment. These results are consistent with previous mutational studies that identified potential interaction residues between DBP-RII and DARC (VanBuskirk *et al.*, 2004, Hans *et al.*, 2005), several of which map to the interaction surfaces identified here.

Residues D264, I265 and T266 are located in the secondary DARC binding interface and directly contact DARC. Mutation of these residues resulted in a loss in RBC binding, demonstrating that the secondary DARC binding interface plays a role in engaging DARC during RBC invasion. Large bulky amino acid changes are required in order to disrupt binding by mutation at the secondary binding site, as is expected from the large contact area created by the additional interfaces in the full complex. The need for large changes is consistent with a lack of an effect on binding when mutations to alanine or conservative changes were introduced at these residues (VanBuskirk *et al.*, 2004, Sampath *et al.*, 2013). Together, these results support the functional role of both DARC binding interfaces.

DARC on erythrocytes is sulfated (Choe *et al.*, 2005). The mutational studies discussed above show that sulfation of DARC on erythrocytes, and the remaining segments of full length DARC, cannot compensate for the loss of the critical binding sites containing L270-K289 and
Q356-K367 that bind DARC 19-30. In particular, recombinant DBP-RII containing Y363A is unable to bind to sulfated DARC in vitro (Hans et al., 2005). This result also demonstrates that sulfation of DARC cannot compensate for the loss of the DARC binding sites induced by the Y363A mutation. Together, the results identify essential binding residues within DBP-RII, consistent with prior studies, which form critical binding sites required for engagement of DARC.

**DBP receptor specificity is manifested through changes in the DARC binding sites**

*P. knowlesi* uses three different DBP homologs to invade human and rhesus macaque RBCs (Chitnis et al., 1994). Only *P. knowlesi* DBPα (PkDBPα) binds DARC, while PkDBPβ and PkDBPγ do not bind human RBCs and recognize a receptor other than DARC (Chitnis et al., 1994). This receptor specificity is likely due to three amino acid changes in the critical DARC binding site of DBP that are changed to non-conservative amino acids in PkDBPβ and/or PkDBPγ (Fig. 6C). The DBP residue with the most extensive DARC contacts is Y363. In both PkDBPβ and PkDBPγ, Y363 is changed to leucine. Mutation of Y363L resulted in a complete loss of binding (Fig. 6D and E) consistent with a role for this residue in receptor specificity. Additionally, R274E and Q356Y in PkDBPβ, as well as Q356D in PkDBPγ, would likely destabilize DARC binding as both R274 and Q356 contact the DARC residues E23 and Y30, respectively. It has been demonstrated that a R274E mutation in DBP-RII completely prevents RBC binding (Batchelor et al., 2011). Mutation of Q356D and Q356Y resulted in a loss in RBC binding, with Q356D having a large effect (Fig. 6D and E). Thus, contacts identified in the DBP:DARC structure provide insight into why PkDBPβ or PkDBPγ do not bind DARC on human RBCs.
4.5 Discussion

Numerous functional and immunological studies have been conducted on the *P. vivax* DBP invasion system since the Duffy antigen was found to be essential to this species in the 1970’s (Miller et al., 1976). Here we have shown using crystallography that DBP-R2I binds DARC forming a heterotrimer and heterotetramer and demonstrated using ITC that the interaction assembles in discrete steps. In addition, we identify DARC residues 19-30 as a critical interaction site for DBP-R2I binding, and thus to tight junction formation during invasion. These studies also identify DBP-R2I residues that directly contact the DARC receptor, including L270-K289, Q356-K367 and F261-T266. The structural and mechanistic basis of Duffy recognition by *P. vivax* provides a context for prior work and may assist with the rational design of therapeutics and vaccines targeting this essential *P. vivax* binding interaction.

The binding pockets in DBP identified here are distinct from a patch of residues previously suggested to engage DARC (Singh et al., 2006) (Fig S4). These residues were proposed based on loss of binding of DBP to DARC upon mutation; however, no data for direct interaction of this patch of residues with DARC was presented. In contrast, the crystal structure of *P. vivax* DBP in complex with DARC demonstrates clear contact points between the two binding partners, and mutagenesis data strongly supports the critical role of the residues identified in the binding pockets.

Phylogenetic studies have identified primate DARC residues under positive selection to block *Plasmodium* infection (Demogines et al., 2012). V25 in DARC is especially polymorphic among primates and under strong positive selection. This is because DBP makes essential hydrophobic packing interactions with V25 (Fig. 7A), disruption of which would strongly destabilize binding. Gorillas, the ancient host of *P. falciparum* (Liu et al., 2010), have a V25A
mutation in DARC. This disrupts hydrophobic interactions, prevents DBP binding (Tourmannille et al., 2005), and serves as an inter-species barrier to *P. vivax* infection.

Sequencing of parasite populations show particular sites of DBP are under strong positive selective pressure to evade the immune response (Xainli et al., 2000, Gosi et al., 2008, Nobrega de Sousa et al., 2011). Many polymorphic DBP residues are located far from the DARC binding sites (Fig. 7B) (VanBuskirk et al., 2004, Hans et al., 2005). The most polymorphic region of DBP, the DEK epitope, forms a ridge directly opposite DARC, flanking the secondary binding interface and homodimer interface. Converting this epitope to small, nonpolar amino acids, focuses the immune response towards cross-specific neutralizing epitopes (Ntumngia et al., 2012). Our results suggest that this hypervariant DEK epitope does not play a direct role in DARC binding. Polymorphisms in the DEK epitope should not affect DBP function, but could interfere with immune recognition of DBP. Antibody recognition of the hypervariant DEK epitope may neutralize *P. vivax* by preventing assembly of the DBP-RII:DARC complex, and thus sterically preventing DBP-RII homodimeric contacts. Polymorphisms are heavily selected for within the DEK epitope suggesting *P. vivax* evades this potentially neutralizing antibody response by antigenic variation within these residues.

The studies presented here define a putative mechanism for known neutralizing epitopes of *P. vivax* DBP-RII and illuminate potential new targets for naturally acquired immunity. Specifically, both DARC helices are oriented in a parallel manner (Fig. 2B and 2D), and DARC itself is a homodimeric GPCR. Because only 30 amino acids, DARC residues G31 to S60, separate the structure from the RBC membrane, the surface of the DBP-RII dimer with DARC Y30 is proximal to the RBC membrane, and the alternate surface faces the *Plasmodium* membrane. Antibody recognition of DBP-RII’s RBC proximal surface prior to DARC binding
neutralizes DBP by sterically preventing DBP-RII from approaching the RBC surface. This model is confirmed by recent work which identified several DBP neutralizing epitopes recognized by human antibodies from individuals living in endemic areas (Chootong et al., 2010) which are located either directly at the DARC binding sites or at DBP-RII’s RBC proximal surface (Fig. 7C). In addition, the most potent known neutralizing epitope for DBP includes much of helix 4 and loop 254-267 (Chootong et al., 2010), which contains the DBP-RII dimerization interface and the DARC binding sites. Disrupting DBP-RII dimerization would both destabilize DARC binding by preventing secondary DBP-RII contacts and destroying interaction contributions due to avidity. Recently, mouse monoclonal antibodies that bound subdomain 3 of DBP-RII also blocked binding to erythrocytes (Siddiqui et al., 2012). Subdomain 3 lies in close proximity to the RBC surface (Fig. 8) and these antibodies may block binding by preventing DBP-RII from approaching the RBC surface and contacting DARC. The identification of two DARC binding sites within DBP-RII and the structural orientation of each molecule provide insight into the mechanisms of antibody inhibition of *P. vivax* RBC invasion. It appears that antibodies targeting DBP-RII are capable of preventing DARC binding by recognizing DBP-RII’s RBC proximal surface, DARC recognition sites, or the homodimeric interface, and may block invasion using other, currently unidentified mechanisms.

The crystallographic and ITC solution studies presented here support a step-wise binding model in which receptor-induced DBP-RII dimerization facilitates formation of a heterotrimer that subsequently recruits a second DARC molecule to form a heterotetramer (Fig. 8). Due to avidity contributions to binding inherent in a two-site mechanism, this heterotetrameric complex may enable the observed tight binding of *P. vivax* to the RBC membrane. Since DBP-RII is monomeric in the absence of DARC (Batchelor et al., 2011), the dimer interface and DARC
binding pockets are exposed and accessible to antibodies prior to DARC engagement. DARC is known to exist as a homodimeric and heterodimeric GPCR (Chakera et al., 2008). The heterotrimer and heterotetramer could represent DBP-RII binding to a DARC heterodimer or homodimer, respectively. The observed transitions may be a selectivity mechanism for DBP-RII to preferentially bind homodimeric DARC while maintaining the ability to bind to a DARC heterodimer. The binding model proposed here is applicable to other DBL domain proteins that may oligomerize upon receptor binding (Tolia et al., 2005, Batchelor et al., 2011, Hodder et al., 2012, Lin et al., 2012, Malpede et al., 2013). Since dimerization is prevalent in receptor signaling, it is plausible that complex assembly initiates a signal through the transmembrane and cytoplasmic domains of DBP to activate pathways of invasion.

Although structure determination of the DBP-RII:DARC complexes allows for visualization and identification of critical contact points, the relevance of each intermediate to complex assembly in solution is not immediately known from the static pictures of binding. To begin to assess the biological role of complex assembly, we utilized ITC to demonstrate that two binding events corresponding to the formation of a heterotrimer and heterotetramer exist in solution. We further tested mutant DBP-RII constructs for RBC binding and demonstrated that these mutations ablate binding to RBCs, supporting the biological role of the DARC contacts identified here as well as the role of the dimer interface. In addition, the large buried surface area for both DARC binding sites and the ITC measurements suggest high affinity interactions. This study thus unambiguously identifies DARC residues 19-30 as a critical binding element that interacts with DBP residues L270-K289, Q356-K367 and F261-T266.

The biphasic profile obtained by ITC is different from studies previously reported where a single binding event with a molar ratio of 1 was observed indicative of the heterotetramer
(Batchelor et al., 2011). This difference is likely due to the buffer conditions used in each case. In prior studies, titrations were performed at a salt concentration of 50 mM while the studies presented here were performed in PBS to examine binding under physiological conditions. These results suggest that observation of the heterotrimer intermediary step by ITC is salt dependent. Never-the-less, the biphasic profile and step wise binding mechanism presented here are representative of the assembly mechanism under physiologically relevant conditions.

In both crystal structures, clear electron density was observed for residues 19-30 of DARC. Mutational studies (VanBuskirk et al., 2004, Hans et al., 2005) and glycan shielding experiments (Sampath et al., 2013) have identified several patches of residues that affect binding of DBP to RBCs, some of which overlap and are consistent with the DARC contacts identified here. There are additional residues outside the DARC binding pockets that when altered reduce binding (VanBuskirk et al., 2004, Hans et al., 2005, Sampath et al., 2013). Therefore, the complete range of interactions between DARC and DBP-R1I1 will likely include additional patches of residues in DBP. Specifically, sulfation of tyrosine 41 and the Fya/Fyb polymorphism have been shown to play a role in binding (Choe et al., 2005, King et al., 2011). However, the specific mechanism by which these changes impact binding is unknown. In addition, an association between the N-terminus and additional extracellular loops of DARC has been suggested to play a role in chemokine binding to DARC (Tournamille et al., 1997). Further studies are necessary to fully understand the role of tyrosine sulfation, the Fya/Fyb polymorphism, and the potential role of additional regions and loops of DARC during *P. vivax* binding and RBC invasion.

The identification of critical DARC binding pockets presented in this study may facilitate the rational design of therapeutics that seek to inhibit RBC binding. Small molecule inhibitors that bind to the DARC binding pocket and prevent DARC engagement could disrupt merozoite
invasion, and thus *P. vivax* growth. Alternately, disruption of complex assembly by small molecules, as has been described for AMA-1:RON-2 (Srinivasan *et al.*, 2013), or antibodies would also provide novel methods for preventing RBC invasion. Since complex assembly is dependent on DARC binding, the most potent disruption of RBC engagement is expected by targeting DARC binding sites. Recent work examining the mechanism of monoclonal antibodies targeting EBL ligands supports the view that targeting receptor binding sites and multimerization interfaces of EBL ligands effectively prevents RBC binding and limits parasite growth *in vitro* (Chen *et al.*, 2013). Additionally, glycan masking experiments with DBP-RII identified the dimer interface and surfaces adjacent to this interface as critical binding sites and targets of an inhibitory antibody response (Sampath *et al.*, 2013). This result supports the importance of identifying and targeting essential functional residues/interfaces of DBP-RII and confirms the biological importance of the contact points identified in this study.

This work also has implications for diagnostics and measures aiming to quantify the immune response to natural infection and in determining the efficacy of vaccine candidates. For a more robust measure of protection, these approaches should quantify the immune response to the functional regions identified here in addition to the response to the entire DBP-RII DBL domain. This study thus expands our understanding of the essential interaction between DBP and DARC and may aid in defining *in vivo* studies that seek to examine the extensive receptor-ligand binding interactions that are essential to RBC invasion by *Plasmodium* species.
4.6 Experimental Procedures

Protein Expression, Purification, and Complex Formation

DBP-RII and DARC were produced as previously described (Batchelor et al., 2011). DARC constructs were expressed in *E. coli* with an N-terminal GB1 tag, followed by a hexahistidine tag and a PreScission Protease cleavage site. Nickel-NTA chromatography followed by PreScission protease treatment and gel filtration resulted in a homogenous sample.

NMR

NMR data were collected at 298 K on a 600 MHz Bruker spectrometer equipped with a triple-resonance room temperature probe and a QCI cryoprobe. Backbone assignments for the non-proline residues in DARC 1-60 were obtained using standard HNCACB, CBCA(CO)NH, HN(CA)CO, and HNCO experiments. Once the DARC1-60 backbone resonances had been assigned, we collected 1H-15N-TROSY spectra of DARC1-60 in the presence of DBP-RII. As DARC residues tightly bound to DBP-RII in a large complex are not visible, peaks which remain in the DBP-RII:DARC 2D 1H-15N TROSY and 3D TROSY triple resonance spectra revealed residues which are not bound by DBP-RII.

Crystallization and Data Collection

Before complex formation, DBP-RII and DARC were purified separately by size-exclusion chromatography to remove any trace aggregates in either sample. The DBP-RII:DARC complexes were prepared by mixing purified DBP-RII and purified DARC in 1:1.2 molar ratio. The DBP-RII:DARC complexes were purified using size-exclusion chromatography in 20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), pH 7.4, and 50 mM sodium chloride. These sample was then concentrated in an Amicon concentrator with a 3-kDa molecular weight cutoff to 20mg ml⁻¹ for crystallization trials.
Native DBP-RII:DARC crystals of both constructs were grown by hanging-drop vapor diffusion by mixing 1 μl of protein solution at 20 mg ml$^{-1}$ and 1 μl of reservoir solution containing 0.1 M HEPES pH 7.4 and 20% (w/v) polyethylene glycol 6000. Crystallization of DBP-RII in complex with DARC 16-43 yielded the heterotrimeric structures, while crystallization of DBP-RII in complex with DARC 14-43 yielded the heterotetrameric crystals. The different crystal forms are not due to the DARC constructs used, rather serendipitous formation of one or either of the stable states upon complex assembly. Crystals were cryoprotected by transfer to reservoir solutions supplemented with glycerol and flash frozen in liquid nitrogen. Data for the heterotetramer was collected at a wavelength of 1.0 Å at beamline 4.2.2 of the Advanced Light Source, Lawrence Berkeley National Laboratory. Data for the heterotrimer was collected at a wavelength of 0.97929 Å at beamline 19-ID of the Advanced Photon Source, Argonne National Laboratory. Data reduction and processing was performed in XDS (Kabsch, 2010). Data collection statistics are shown in Table 1.

**Structure Solution**

Both structures were solved by molecular replacement using DBP-RII apo-structure (Batchelor et al., 2011) leading to a starting model with Rwork/Rfree of 37.3%/38.1% for the heterotrimer and an Rwork/Rfree of 30.0%/30.1% for the heterotetramer. NCS restraints were not imposed on the two copies of DBP-RII during refinement as it was clear from electron density maps they were not identical. Subsequent automated rebuilding in PHENIX AutoBuild (Adams et al., 2010), manual rebuilding in COOT0.7 (Emsley et al., 2010) and refinement in PHENIX1.7.3 (Adams et al., 2010) lead to a final model with Rwork/Rfree of 16.62%/20.47% for the heterotrimer and 18.29%/23.28% for the heterotetramer (Table 1). These low R factors combined with the good Ramachandran plot statistics analyzed by MolProbity
(Davis et al., 2007) indicated that structure refinement was complete. Residue distributions in the Ramachandran plot for the heterotrimer were 98.42 % allowed, 1.58 % additionally allowed and 0 % disallowed. Residue distributions in the Ramachandran plot for the heterotetramer were 95.72 % allowed, 4.28 % additionally allowed and 0 % disallowed. The atomic coordinates and structure factors for the structure have been deposited in the protein data bank with accession numbers 4NUV and 4NUU.

**Isothermal Titration Calorimetry**

DBP-RII was prepared as described above, with the addition of an ion-exchange chromatography step prior to ITC measurements. DBP-RII and DARC1-60 were exchanged into PBS to ensure measurements were made under physiological conditions. ITC experiments were carried out at 10°C using a VP-ITC instrument (MicroCal). DARC1-60 at 1.3 mM was titrated into 1.4 mL of 130 μM DBP-RII. Traces were analyzed using Origin Version 5.0 (MicroCal). Stoichiometry and binding constants were calculated by fitting the integrated data to an independent two-site binding model. Protein concentrations were determined by absorbance measurements under denaturing conditions (6 M guanidinium hydrochloride, 10 mM dithiothreitol).

**Functional Studies**

DBP-RII with a C-terminally fused green fluorescent protein (GFP) was cloned into plasmid pRE4 for surface expression in mammalian cells. Single-amino-acid mutations were introduced in DBP-RII using the QuickChange method (Stratagene). Fresh monolayers of HEK293T cells were cultured in 3.5-cm-diameter wells and transfected with 2 μg ml⁻¹ DNA in polyethyleneimine. The binding assay was performed 20 h after transfection. Anonymized human RBCs (ZenBio) were added to each well in a 10% suspension, incubated at 37 °C for 1 h,
and washed three times with PBS. Binding was quantified by counting rosettes observed over ten fields of view at x200 magnification. Transfected HEK 293T cells with five or more attached RBCs were defined as positive rosettes. In each experiment, three wells of HEK 293T cells were transfected for each mutation. Cell counting was performed using ImageJ (NIH) on randomized images. Three fields of view from ten independent transfections (final n = 30) were counted for each sample (wild type or mutant). Significance was tested by a paired two-tailed Student's t-test as the data were normally distributed and had large sample sizes (n = 30).
4.7 Figures and Tables

**Figure 1:** Residues 14-43 of DARC contain the minimal binding region. $^1$H-$^{15}$N-TROSY spectra of unbound DARC 1-60 (black) overlaid on $^1$H-$^{15}$N-TROSY spectra of DARC 1-60 in the presence of excess unlabelled DBP-RII (red). Sequence assignments are shown for the unbound DARC $^1$H-$^{15}$N-TROSY spectra. Peaks still visible in the presence of DBP-RII (red) are at DARC 1-60's N- and C- termini. Residues that disappear in the presence of DBP-RII are in the center of DARC and delineate the binding region.
Figure 2: Crystal Structure of the DBP-RII:DARC heterotrimer and heterotetramer. Overview of (A) and DBP-RII:DARC heterotrimer and (B) the DBP-RII:DARC heterotetramer. Rotated views, (C) and (D), show DARC helices are oriented in parallel in the heterotetramer. DBP-RII monomers are in yellow and green. DARC monomers are in purple and blue.
Figure 3: Isothermal titration calorimetry reveals step-wise binding of DARC to DBP-RII in solution. (A) A biphasic binding profile is observed indicating the formation of the heterotrimer at a molar ratio of 0.5 (n = 0.45±0.03, K_{d1} = 2151±352 nM, ΔH_{1} = -2725±153 cal/mol) and heterotetramer at a molar ratio of 1 (n = 0.48±0.04, K_{d2} = 56.8±50.4 nM, ΔH_{1} = -3339±57 cal/mol). Molar ratios are expressed as monomers of DBP-RII. Open circles denote unbound DBP, closed circles denote bound DBP. Titration of (B) PBS into DBP and (C) DARC into PBS reveals no observable profiles demonstrating the biphasic profile is due to DARC binding to DBP. In all cases, the top panel contains raw binding data, and the bottom panel changes in enthalpy associated with binding.
Figure 4: Binding interfaces of the DBP-RII:DARC heterotrimer. (A) Global view of the DBP-RII:DARC heterotrimer, showing (B) DARC monomer A interactions and (C) the DBP-RII homodimeric interface. DARC monomer A is in purple, DBP-RII monomer 1 is in green and DBP-RII monomer 2 is in yellow. Residue numbers are labeled and DARC residue labels are underlined.
Figure 5: Binding interfaces of the DBP-RII:DARC heterotetramer. (A) Global view of the DBP-RII:DARC heterotetramer, showing (B) the DBP-RII homodimeric interface, (C) DARC monomer A interactions, and (D) DARC monomer B interactions. DARC monomer A is in purple, DARC monomer B is in blue, DBP-RII monomer 1 is in green and DBP-RII monomer 2 is in yellow. Residue numbers are labeled and DARC residue labels are underlined.
Figure 6: The structural studies define red blood cell binding. (A) Adherent HEK293 cells in grey bind to darker, smaller red blood cells when transfected with a DBP-RII surface expression plasmid with a GFP marker. Red blood cell rosetting images for DBP-RII mutants, showing bright field (left), GFP (center), and merged images (right). (B) Percentage of cells expressing point mutants which bind red blood cells, relative to wildtype, shown with standard error. (C) The major DBP-RII:DARC residues identified in the crystal structures are indicated by red dots. Non-conservative *P. knowlesi* mutations at critical DBP-RII:DARC contact residues 274, 356, and 363 suggest why PkDBPα but not PkDBPβ or PkDBPγ bind DARC. (D) Red blood cell rosetting images for DBP-RII receptor specificity mutants, showing bright field (left), GFP (center), and merged images (right). (E) Percentage of cells expressing receptor specificity point mutants which bind red blood cells, relative to wildtype, shown with standard error.
Figure 7: Mapping polymorphic residues and inhibitory epitopes reveals targets of selective pressure. DBP-RII molecules are in green and yellow. DARC molecules are in purple and blue. DARC residue labels are underlined. (A) Nonsynonymous DARC polymorphisms in primates, blue, which make critical contacts with DBP-RII provide a mechanism for inter-species transmission barriers. (B) Polymorphic DBP residues, in blue, are spread throughout the molecule. The most polymorphic region of DBP is the “DEK epitope” opposite the DARC14-43 binding site. (C) Inhibitory epitopes, in red and brown, map to the heterotetramer interface, DARC binding pockets and RBC proximal face of DBP-RII.
Figure 8: A model for attachment during invasion. An initial binding event is followed by receptor-induced dimerization, as in the DBP-RII:DARC heterotrimer. This brings a second DBP-RII molecule in close proximity to a second DARC ectodomain in the DARC homodimer. A second binding event creates the DBP-RII:DARC heterotetramer. DBP-RII molecules are in green and yellow and DARC19-30 molecules are in purple and blue. The DARC homodimer is represented by a homology model. A schematic for the stepwise assembly is show at the bottom. Closed circle – bound DBP-RII, open circle – unbound DBP-RII.
**Figure S1:** DARC residues 19-30 are contacted by DBP-RII. DARC19-30 binds to a positively charged groove at the DBP-RII dimer interface of both (A) the DBP-RII:DARC heterotrimer and (B) the DBP-RII:DARC heterotetramer. Electrostatic potential is shown from -7.5 to 7.5 kT/e with positive potential in blue and negative potential in red. 2fo-fc electron density maps, contoured at 1σ clearly show the presence of (C) a single DARC19-30 in the heterotrimer and both (D) DARC19-30A and (E) DARC19-30B in the two DBP-RII binding sites of the heterotetramer. DARC monomers are in purple and blue and DBP-RII monomers are in green and yellow.
Figure S2: Upon receptor binding, new regions of DBP-RII become structured, while preexisting structural regions undergo no major conformational changes. During the transition from the heterotrimeric to heterotetrameric complex, a change in the overall architecture of the DBP-RII dimer is observed. In (A-C) the DARC-bound DBP-RII heterotetramer is green and yellow, the DARC-bound DBP-RII heterotrimer is light green and light yellow, and unbound DBP-RII is dark green and orange. Structural transitions in each case are designated with an arrow as well as with the distance of the structural shift. (A) A translation covering 12 Å along helix 4 defines the difference between the heterotrimeric structure and a prior structure of DBP-RII in the absence of receptor. (B) A translation covering 12 Å across helix 4 is the difference between the heterotrimeric structure and the heterotetrameric structure. (C) A translation covering 23 Å along helix 4 is the difference between the heterotetrameric structure and DBP-RII in the absence of receptor, which defines the full shift following binding of both DARC molecules. (D-G) Alignments of the individual monomers of the DBP-RII:DARC heterotetramer and unbound DBP-RII. (D) Monomer A of the heterotetramer (green) with monomer A unbound (dark green), (E) monomer A of the heterotetramer (green) with monomer B (orange) unbound, (F) monomer B of the heterotetramer (yellow) with monomer A unbound (dark green), (G) monomer B (yellow) of the heterotetramer with monomer B unbound (orange).
**Figure S3:** The Sulfotyrosine Binding Site. DBP-RII molecules are in green and yellow. DARC molecules are in purple and blue. (A) Phosphate or selenate in the apo DBP-RII structure occupy the same position as (B) DARC Y30, defining the sulfotyrosine binding pocket.
**Figure S4:** The DARC binding pockets are distinct from residues previously suggested to bind DARC from mutagenesis studies. DBP-RII monomers are in yellow and green. DARC monomers are in purple and blue. Residues previously suggested (Singh *et al.*, 2006) to contact DARC are in black.
<table>
<thead>
<tr>
<th>Table 1: Data Collection and Refinement Statistics</th>
<th>DBP-RII:DARC heterotrimer</th>
<th>DBP-RII:DARC heterotetramer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data collection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P2(1)</td>
<td>P1</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a, b, c$ (Å)</td>
<td>59.59, 66.99, 97.92</td>
<td>37.39, 59.47, 91.67</td>
</tr>
<tr>
<td>$\alpha, \beta, \gamma$ (°)</td>
<td>90, 102.112, 90</td>
<td>103.143, 91.098, 100.241</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>20-1.95 (2.0-1.95)</td>
<td>20-2.6 (2.7-2.6)</td>
</tr>
<tr>
<td>$R_{sym}$</td>
<td>.098 (1.01)</td>
<td>0.052 (0.538)</td>
</tr>
<tr>
<td>$I / \sigma I$</td>
<td>14.54 (2.11)</td>
<td>12.32 (1.69)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.8 (99.7)</td>
<td>95.2 (93.1)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>7.96 (7.96)</td>
<td>2.4 (2.2)</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>20-1.95</td>
<td>20-2.6 Å</td>
</tr>
<tr>
<td>No. Reflections</td>
<td>58,872</td>
<td>22,121</td>
</tr>
<tr>
<td>$R_{work} / R_{free}$</td>
<td>16.62/20.47</td>
<td>18.29/23.28</td>
</tr>
<tr>
<td>No. Atoms (Non-Hydrogen)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>5,422</td>
<td>5,607</td>
</tr>
<tr>
<td>Ligand Organic</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>Water</td>
<td>462</td>
<td>52</td>
</tr>
<tr>
<td>B-factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>38.86</td>
<td>70.62</td>
</tr>
<tr>
<td>Ligand Organic</td>
<td></td>
<td>74.64</td>
</tr>
<tr>
<td>Water</td>
<td>39.44</td>
<td>52.38</td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.005</td>
<td>0.002</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>0.832</td>
<td>0.543</td>
</tr>
</tbody>
</table>

*Values in parentheses are for highest-resolution shell.

Data were collected on a single crystal for each dataset.
Table 2: Heterotrimer interface residues determined by PDBePISA (Krissinel et al., 2007): all residues in the interface are listed sequentially and do not indicate interacting pairs

<table>
<thead>
<tr>
<th>DARC binding site</th>
<th>Dimer Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary DARC binding interface (DARC A to DBP 1)</strong></td>
<td><strong>Secondary DARC binding interface (DARC A to DBP2)</strong></td>
</tr>
<tr>
<td>DARC A</td>
<td>DBP 1</td>
</tr>
<tr>
<td>GLN19</td>
<td>LEU270</td>
</tr>
<tr>
<td>LEU20</td>
<td>LYS273</td>
</tr>
<tr>
<td>ASP21</td>
<td>ARG274</td>
</tr>
<tr>
<td>PHE22</td>
<td>ILE277</td>
</tr>
<tr>
<td>GLU23</td>
<td>TYR278</td>
</tr>
<tr>
<td>ASP24</td>
<td>ALA281</td>
</tr>
<tr>
<td>VAL25</td>
<td>VAL282</td>
</tr>
<tr>
<td>TRP26</td>
<td>ASP285</td>
</tr>
<tr>
<td>SER28</td>
<td>GLN356</td>
</tr>
<tr>
<td>SER29</td>
<td>THR359</td>
</tr>
<tr>
<td>TYR30</td>
<td>ALA360</td>
</tr>
<tr>
<td></td>
<td>TYR363</td>
</tr>
<tr>
<td></td>
<td>SER364</td>
</tr>
<tr>
<td></td>
<td>LYS366</td>
</tr>
</tbody>
</table>
Table 3: Heterotetramer interface residues determined by PDBePISA (Krissinel et al., 2007): all residues in the interface are listed sequentially and do not indicate interacting pairs.

<table>
<thead>
<tr>
<th></th>
<th>DARC binding site 1</th>
<th>DARC binding site 2</th>
<th>Dimer Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Primary DARC binding interface (DARC A to DBP 1)</td>
<td>Secondary DARC binding interface (DARC A to DBP 2)</td>
<td>Primary DARC binding interface (DARC B to DBP 1)</td>
</tr>
<tr>
<td>DARC A</td>
<td>DBP 1</td>
<td>DARC A</td>
<td>DBP 2</td>
</tr>
<tr>
<td>GLN19</td>
<td>GLU249</td>
<td>GLU23</td>
<td>PHE261</td>
</tr>
<tr>
<td>LEU20</td>
<td>LEU270</td>
<td>LEU20</td>
<td>LYS273</td>
</tr>
<tr>
<td>ASP21</td>
<td>LYS273</td>
<td>ASP21</td>
<td>ARG274</td>
</tr>
<tr>
<td>PHE22</td>
<td>ARG274</td>
<td>PHE22</td>
<td>LYS275</td>
</tr>
<tr>
<td>GLU23</td>
<td>LYS275</td>
<td>GLU23</td>
<td>TRP26</td>
</tr>
<tr>
<td>ASP24</td>
<td>ILE277</td>
<td>ASP24</td>
<td>HIS262</td>
</tr>
<tr>
<td>TRP26</td>
<td>ALA281</td>
<td>TRP26</td>
<td>ASN27</td>
</tr>
<tr>
<td>SER28</td>
<td>VAL282</td>
<td>SER28</td>
<td>ARG263</td>
</tr>
<tr>
<td>SER29</td>
<td>ASP285</td>
<td>SER29</td>
<td>THR359</td>
</tr>
<tr>
<td>TYR30</td>
<td>LYS289</td>
<td>TYR30</td>
<td>ALA360</td>
</tr>
<tr>
<td>GLN356</td>
<td>THR359</td>
<td>GLN356</td>
<td>TYR363</td>
</tr>
<tr>
<td>THR359</td>
<td>ALA360</td>
<td>THR359</td>
<td>SER364</td>
</tr>
<tr>
<td>ALA360</td>
<td>TYR363</td>
<td>ALA360</td>
<td>LYS366</td>
</tr>
<tr>
<td>TYR363</td>
<td>SER364</td>
<td>TYR363</td>
<td>LYS367</td>
</tr>
<tr>
<td>SER364</td>
<td>LYS367</td>
<td>SER364</td>
<td>LYS367</td>
</tr>
</tbody>
</table>
4.8 References


Chapter 5

Conclusions and Future Directions
5.1 Introduction

The hypothesis driving this thesis work is that individual EBL ligands utilize distinct receptor binding interfaces and mechanisms to engage their unique receptors on the red blood cell. I presented data that supports the presence of unique binding mechanisms for EBL ligands and propose that these distinct elements define receptor specificity for individual invasion pathways. Examining all members of this family is crucial as the proposed redundancy in function provides separate routes for RBC entry, effectively protecting the parasite from immunity and receptor polymorphism.

The work presented here elucidated key binding interface and mechanistic differences between two unique EBL members, PfEBA-140 and PvDBP, and allowed for comparison with a third characterized family member, PfEBA-175. Initial efforts elucidated the oligomeric state and sialic acid binding properties of PfEBA-140, and defined determinants of receptor specificity within the \textit{P. falciparum} EBL family. The work expanded to PvDBP with the goal of identifying unique aspects of invasion between the \textit{Plasmodium} species. Work on PvDBP focused on the physical receptor engagement mechanism, and demonstrated a two-step, putatively cooperative process for receptor binding.

The long term objective of this work is to exploit the defined binding elements and mechanisms utilized by EBL members in an effort to enhance the immune response to these invasion pathways. Current immunization methods utilizing full length RII constructs provide the immune system with access to decoy and non-inhibitory epitopes, potentially reducing the efficacy of the response. Thus, focusing the immune system on functional binding elements identified by crystallographic, biochemical, and mutational studies may limit the elicitation of non-inhibitory antibodies and enhance the inhibitory antibody response. The combination of
structural, biophysical, and cellular interaction techniques discussed in this work represent an experimental system that allows for the delineation of atomic and molecular details defining the interface between *Plasmodium* parasites and their host, and may ultimately aid in developing therapeutics and vaccines targeting malaria. Below I discuss the biology and evolution of *Plasmodium* RBC invasion mechanisms, the proposal to specifically target biochemical invasion interfaces for vaccine development, and potential future directions.
5.2 Receptor engagement mechanisms of the EBL family

The EBL ligands each bind to unique red blood cell receptors utilizing either a single or dual-DBL domain mechanism. This protein architecture differs between *P. falciparum* and *P. vivax*. Insight into mechanisms and binding elements provided in this work are unique for the most well characterized EBL members, including PvDBP, PfEBA-140, and PfEBA-175. Mechanisms and binding interfaces for other EBL ligands have not been examined in detail, and thus will not be discussed in the context of this thesis research.

The proposed binding mechanisms for EBL ligands can be separated into three groups based on the number of DBL domains in RII and the oligomeric state observed during receptor engagement. Proposed mechanisms include: 1) a single DBL domain RII with a dimeric receptor interaction; 2) a dual DBL domain RII with a dimeric receptor interaction; and 3) a dual DBL domain RII with a monomeric receptor interaction (Figure 1). It is plausible that additional biophysical invasion mechanisms exist for the EBL family.

Understanding the mechanisms of binding for the DBL domain is crucial towards defining the full invasion repertoire of the parasite. The presence of the dual DBL domain RII in PfEBA-175 provides both unique receptor binding interfaces and multimeric interfaces within the DBL domain compared with PvDBP, which contains only one DBL domain in RII. However, PfEBA-175 and PvDBP are mechanistically similar in that they appear to dimerize while binding their cellular receptors. In contrast to the multimeric mechanism, all current evidence points to a monomeric invasion process for PfEBA-140. This evidence includes SAXS solution studies, albeit in the absence of receptor, and the absence of biologically relevant multimeric interfaces in the crystal structure. PfEBA-140 also exhibits a distinct sugar binding and receptor binding
interface on the DBL domain. In summary, each EBL ligand is characterized by distinct binding elements on the conserved DBL domain.

In addition to distinct biophysical mechanisms, it is highly likely that receptor binding interfaces on the DBL domains of uncharacterized EBL ligands and other adhesive proteins in *P. falciparum* are unique. DBL domains of other adhesive proteins, specifically PfEMP1, do in fact contain unique binding interfaces on the DBL domain (Higgins, 2008, Khunrae *et al.*, 2009, Vigan-Womas *et al.*, 2012). Mechanisms for adhesion of these proteins have not been examined in detail, in part due to the complicated, modular nature of the PfEMP1 family, but it is proposed that higher order multimers are involved (Kraemer *et al.*, 2006, Batchelor *et al.*, 2011). The DBL domain is capable of binding diverse protein receptors, and also interacts with sugar molecules, as discussed for the cases of PfEBA-175 and PfEBA-140 (Tolia *et al.*, 2005). It is this plasticity of receptor binding interfaces and mechanistic capabilities that makes the DBL domain such useful tool for the parasite’s cellular adhesion requirements.

In addition to oligomeric state, receptor binding by the DBL domains of PfEBA-140 and PvDBP induces shifts in physical structure. These shifts likely provide access to the highest affinity receptor binding and oligomeric interfaces. PfEBA-140 exhibits a relatively small shift of a helix near its sialic acid binding site in the F2 DBL domain, possibly to provide access to the protein backbone of GPC. The RBC binding studies presented here suggest that the F2 DBL domain is not essential for sugar binding, but instead engages the critical protein backbone of GPC. It is plausible that PfEBA-140 will exhibit larger shifts in the presence of full length receptor (which includes higher order glycosylation and the protein backbone), and future studies may examine these structural shifts and the complete binding interface. The clear next step is to
identify the necessary interaction surface on GPC and determine the mechanistic role of structural shifts in DBL domains.

In contrast to the small shift observed for PfEBA-140, the step-wise process of PvDBP engaging DARC involves several large shifts in the dimeric interface, which may enhance dimerization and thus the avidity of the interaction. PvDBP is similar to PfEBA-140 in that it appears that the minimal binding portion of DARC is not the only element of the protein involved in binding. The portion identified in this work is the main binding interface; however, recent work in the lab suggests that regions outside of this interface may enhance dimerization through transient, lower affinity contacts. As the PvDBP:DARC binding interface identified here required truncation of the extracellular portion of DARC, accessory contacts could not be identified and will require further study to be defined in full.

In addition to a putative role in driving receptor specificity, invasion ligand multimerization is proposed to activate downstream parasite effectors that initiate the merozoite’s downstream invasion machinery (Singh et al., 2010). Biological signaling systems often utilize dimerization to activate downstream pathways (Klemm et al., 1998). It is plausible that dimerization of the extracellular portion of EBL ligands affects the cytoplasmic region, which has no identified function but is properly positioned to interact with internal signaling molecules. Examination of the interaction between EBL ligands with their receptors demonstrated altered calcium levels in merozoites, prompting rhoptry organelle release, supporting a role in signaling (Singh et al., 2010, Harvey et al., 2012, Weiss et al., 2015). Signaling proteins and a pathway for this mechanism have not yet been identified, but are critical in clarifying the molecular network that functions during invasion and the specific role of EBL proteins.
Multimerization of EBL ligands may play a role in downstream signaling during invasion, however, this mechanism cannot explain the role of EBL ligands with proposed monomeric receptor interactions such as PfEBA-140. Future work must completely define the oligomeric state of PfEBA-140 during engagement of GPC, as previous studies were performed in the absence of receptor or only in the presence of sialic acid, a key component of the receptor but not the receptor itself. Additionally, it will be valuable to examine the specifics of downstream signaling activation during EBL receptor engagement via this monomeric receptor-ligand interaction. It is plausible that a unique signaling mechanism is involved for the PfEBA-140 pathway and other monomeric invasion routes, or that signaling by this pathway is not required to complete the invasion process and is prompted by other molecular interactions.

PfEBA-140 and PvDBP represent two unique members of the EBL family in terms of their apparent invasion mechanisms and binding interfaces discussed above. I have proposed that novel biochemical interfaces as well as biophysical mechanisms determine receptor specificity within this parasite invasion ligand family, and these novelties allow the parasite to utilize distinct receptor routes into the RBC. It is unclear why the mechanism of invasion would be different between EBL members, and the evolution of these molecular invasion properties has not been examined in detail. Furthermore, a proposal describing the origin of multimeric assembly for EBL ligands suggests that ligand oligomeric state is at least partially dependent on receptor oligomeric state. In fact, the dimeric nature of PfEBA-175 and PvDBP fit with the observed native oligomeric state for their respective receptors (Chasis et al., 1992, Lemmon et al., 1992, Treutlein et al., 1992). Additionally, PfEBA-140, appearing to bind as a monomer, fits with the proposed monomeric state of its receptor GPC (Chasis et al., 1992). These observations fit into a model of human and parasite co-evolution over 100,000 – 180,000 years proposed
several years ago and suggest that the parasite has evolved a number of immune evasion mechanisms (Mu et al., 2002, Crompton et al., 2014). Future work should examine the mechanisms for other EBL ligands and identify the native oligomeric state of corresponding receptors. These studies will aid in defining the evolution of unique RBC invasion mechanisms and specific biochemical and biophysical determinants of receptor specificity.
5.3 Targeting functional elements of DBL domains for vaccination

A critical aspect of disrupting merozoite entry into the RBC will be direct inhibition of functional receptor interaction interfaces and blockage of biophysical invasion elements, such as those identified in this work. It is proposed that one mechanism of antibodies targeting highly abundant surface proteins on merozoites is opsonization and/or the recruitment of complement (Osier et al., 2014, Boyle et al., 2015). These mechanisms contrast the mechanistic and physical interaction block demonstrated for invasion ligands expressed only during the invasion process (Chen et al., 2013, Wright et al., 2014). Thus, invasion ligands that are exposed at a specific time during invasion must be recognized by antibodies and also physically disrupt contact between the parasite and the RBC. A focus on targeting functional elements of these invasion proteins is necessitated by the presence of decoy epitopes that may focus the B-cell response and diminish the production of highly inhibitory antibodies (Godson et al., 1983, Godson et al., 1984, Chen et al., 2013).

Efforts in defining a vaccine that seeks to block RBC invasion have historically proposed utilizing multiple antigens to combat parasite immune evasion mechanisms and apparent redundancy in invasion pathways (Pandey et al., 2013). Targeting a single invasion ligand is thought to be ineffective due to invasion pathway switching or the presence of a population of merozoites utilizing a different ligand will overcome the developed immunity. Despite these issues, studies in which Aotus monkeys were immunized with recombinant PfEBA-175 RII protein or by plasmid DNA were able to demonstrate the induction of immunity to lab strains of \textit{P. falciparum}, which supports the efficacy of the response to EBL proteins (Jones et al., 2001). This study was performed with full length RII, thus the immune system had access to variant and decoy epitopes, which may have diminished response efficacy. It has been observed that humans
in endemic areas do in fact develop a B-cell response to EBL ligands, specifically RII, further supporting the role of these proteins in responding to the parasite (Grillot et al., 1990, Ford et al., 2005, Ford et al., 2007, Ismail et al., 2014). However, these studies do not examine specific epitopes or demonstrate a conclusive role in protection, thus it is difficult to conclude what function anti-EBL antibodies play in clearing the parasite and challenging to define immune evasion mechanisms for invasion proteins.

The studies presented in this thesis work identified the binding and dimerization interfaces of PvDBP as the target of inhibitory antibodies, and illuminated regions outside of the functional interfaces as targets of non-inhibitory antibodies. Examination of another merozoite surface protein involved during invasion illuminated an immune escape mechanism based on antigenic diversity (Dutta et al., 2007). Polymorphism in regions outside of functional interfaces is seen in PvDBP as well as other EBL ligands. These observations support an immune escape mechanism based on amino acid changes as non-inhibitory antibodies appear to target these polymorphic regions. It thus seems evident that conserved mechanistic and binding interfaces represent ideal elements for rational immunization strategies.

Studies examining the T-cell response to a merozoite surface invasion protein demonstrated that a subset of cells targeting immunodominant T-cell epitopes are cytolytic while a second subset has no cytolytic activity (Hafalla et al., 2013). This result suggests that while a portion of targeted epitopes develop a productive response, immunodominant regions can also induce a non-inhibitory response. This disparity may also apply to epitopes of the B-cell system, a proposal supported by polymorphism in surface adhesins and the identification of decoy epitopes in invasion ligands (Dutta et al., 2007, Chen et al., 2013).
Based on the identification of the PvDBP receptor binding interface and inhibitory epitopes described here, it is expected that the sialic acid binding pockets identified on one face of PfEBA-140 RII will be heavily targeted by inhibitory antibodies. In contrast, the opposite face of the protein likely serves as a decoy region with immunodominance, similar to the regions outside of the dimerization and DARC binding interface of PvDBP. Future studies should examine inhibitory and non-inhibitory epitopes to define the overall response to RII. In addition, studies should expand into other segments of the EBL proteins. C-terminal extracellular elements of EBL ligands may provide greater capacity to dimerize around a receptor or induce downstream signaling (Withers-Martinez et al., 2008, Paing et al., 2014). A large portion of the EBL protein architecture is uncharacterized in terms of function, and could be examined for both functional and antigenic capacity (Wanaguru et al., 2013). It is plausible that large portions of the EBL protein presented to the immune system limit the functional response to the EBL invasion pathways. In contrast to all other invasion proteins of the EBL and RH families, RH5 is unique in that it is the smallest in terms of amino acid length and molecular weight. RH5 seems to induce some of the most potent inhibitory antibodies that inhibit numerous lab and wild P. falciparum strains, suggesting a strong response to conserved functional elements that is parasite strain-transcending. It is plausible that the extended protein architecture of most invasion ligands of the EBL and RH families evolved to provide each ligand with a level of immune protection that has not yet developed for RH5. Future studies may examine the evolution of the size, shape, and overall role of the portions of invasion proteins outside of the minimal receptor binding domains of RII for the EBL family and the receptor binding element for the RH family.
5.4 EBL ligands in a multi-component vaccine

Recent studies identified RH5 as a merozoite ligand that is potentially essential to the invasion process due to its potent inhibition by antibodies, as discussed above. RH5 is also postulated to be a potentially critical contributor to human immunity and thus an essential component in a vaccine. However, high antibody titers are typically necessary for productive immunity to merozoite ligands, and this is the case for RH5. In fact, even at 10mg/mL of polyclonal IgG, in vitro parasite growth inhibition reaches less than 80%, suggesting there are still populations of parasites productively invading RBCs (Chen et al., 2014, Wright et al., 2014). Immunization with RH5 alone has proven effective in Aotus monkeys; however, these studies were small and involve only one re-infection with various parasite lab strains (Douglas et al., 2015). Trials in large human populations are currently underway and will provide greater insight into RH5 function and antigenicity. However, taking into consideration the inability to completely eliminate parasite growth in vitro at high antibody concentrations and lack of evidence for rapid, long-lasting immunity to RH5, it is critical to identify ideal partners in a multi-component vaccine (Williams et al., 2012).

The key difference between the small studies performed for RH5 and examination of protection in endemic areas is the parasite’s tremendous genetic diversity and the prospect of consistent re-infection. Focusing on a single invasion ligand for immunity thus provides the parasite with a quick genetic escape from the vaccine. A study of a less virulent Plasmodium species in a mouse model of infection demonstrated the evolution of more virulent parasites due to immunization with a single merozoite invasion ligand, supporting the necessity of a multi-component vaccine (Barclay et al., 2012). Furthermore, the polymorphic nature of most invasion
ligands and the identification of numerous non-inhibitory antibodies suggest that part of the merozoite’s immune evasion mechanism involves diminishing the productive response.

Similar to antibody epitopes identified for PvDBP and PfEBA-175, mapping of RH5 inhibitory epitopes illuminated functional interfaces as targeted regions, likely causing a block in the mechanism of binding. Studies examining the epitopes of non-inhibitory antibodies are not readily available, although it was demonstrated that non-inhibitory antibodies do not bind to RH5, suggesting that a subset of induced antibodies are ineffective and do not even maintain the ability to bind the antigen (Wright et al., 2014). This result supports the presence of a parasite immune evasion mechanism that provides the B-cell system with access to immunodominant but non-inhibitory epitopes, effectively limiting productive inhibition.

In addition to the functional interfaces of adhesions assessed in this study, other RBC invasion ligands and blood-stage adhesion proteins containing the DBL domain are being targeted and demonstrate the value of identifying specific functional elements for immunization. In the case of PfEMP1 CIDR domains, a highly alpha helical domain that functions similarly to the DBL domain, conserved structural elements allow the parasite to increase the level of polymorphism without losing receptor binding function (Lau et al., 2015). Perhaps a common thread exists in terms of antigenic potential of the DBL domain. If a common structural or sequence theme among non-inhibitory, immunodominant elements were identified, it would open avenues to understand the complete immune evasion mechanisms of the parasite and allow for rational design of epitopes that may generate greater specificity in the B-cell response.

The thesis research presented here identified specific regions of proteins for targeted immunization in a multi-component vaccine that disrupts RBC invasion, focusing specifically on the DBL domain. Epitopes can be targeted in several ways, specifically by mutating immune-
dominant non-inhibitory epitopes and by shrouding non-inhibitory epitopes with glycosylation (Ntumngia et al., 2012, Sampath et al., 2013). Similar approaches effectively exploited broadly-neutralizing epitopes to viral antigens, supporting the efficacy of these techniques (Corti et al., 2013). It has been established that antibodies are crucial during the response to the parasite, as demonstrated by the alleviation of disease through transfer of purified IgG from immune individuals to those with active malaria infection (Cohen et al., 1961). However, it is still not clear what proteins are the most effective targets or what antibody mechanisms provide the greatest and longest lasting levels of protection. Repeated infection is proposed to elicit protective antibodies targeting PfEMP1, supporting the efficacy of the immune response to surface exposed adhesins containing the DBL fold (Bull et al., 1998). The question still remains as to how to best formulate a vaccine and how to present the antigen(s) to the human immune system. The work presented here advances our understanding of the Plasmodium DBL domain and provides a greater capability to target specific functional and structural elements of two EBL invasion ligands for rational vaccination efforts.
5.5 Figures

**Figure 1.** Summary of proposed invasion mechanisms for *Plasmodium* EBL family members. PfEBA-140 (left) utilizes the dual-DBL domain monomeric engagement mechanism. PvDBP (middle) engages its receptor in a two-step mechanism using a single DBL domain to form dimer. PfEBA-175 (right) also dimerizes around its receptor, but does so with a dual DBL domain RII.
5.6 References


